



(12)发明专利申请

(10)申请公布号 CN 106462617 A

(43)申请公布日 2017.02.22

(21)申请号 201580029053.9

P·M·利斯特 R·达萨里

(22)申请日 2015.03.27

(74)专利代理机构 北京市金杜律师事务所

(30)优先权数据

62/019,312 2014.06.30 US

11256

14/498,503 2014.09.26 US

代理人 王茂华 张宁

(85)PCT国际申请进入国家阶段日

2016.11.30

(51)Int.Cl.

G06F 17/30(2006.01)

(86)PCT国际申请的申请数据

G09C 1/00(2006.01)

PCT/US2015/023089 2015.03.27

G10L 15/22(2006.01)

(87)PCT国际申请的公布数据

H04N 21/41(2011.01)

W02016/003509 EN 2016.01.07

H04N 21/422(2011.01)

H04N 21/431(2011.01)

(71)申请人 苹果公司

地址 美国加利福尼亚州

(72)发明人 M·范奥斯 H·J·萨德勒

L·T·拿波利塔诺 J·H·拉塞尔

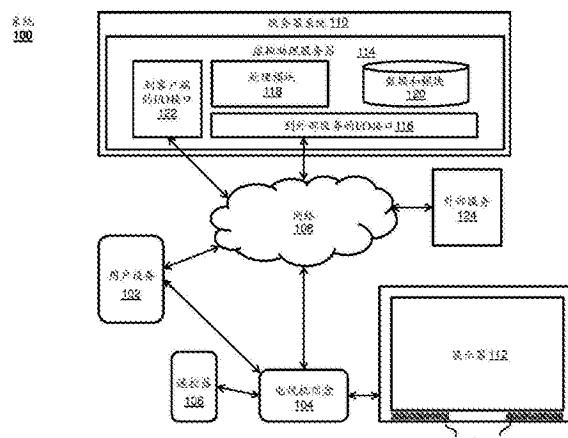
权利要求书8页 说明书41页 附图42页

(54)发明名称

用于电视机用户交互的智能自动化助理

(57)摘要

本发明公开了一种用于使用虚拟助理来控制电视机用户交互的系统和方法。该虚拟助理可与电视机顶盒进行交互，以控制在电视机上所显示的内容。可从具有麦克风的设备接收针对虚拟助理的语音输入。可从语音输入来确定用户意图，并且虚拟助理可根据用户意图来执行任务，包括使得在电视机上回放媒体。虚拟助理交互可被显示在界面中的电视机上，该界面扩展或收缩以占据最小量的空间，同时传送期望的信息。可使用与多个显示器相关联的多个设备从语音输入确定用户意图，以及向用户传送信息。在一些示例中，可基于在显示器上所显示的媒体内容来向用户提供虚拟助理查询建议。



1. 一种用于使用虚拟助理来控制电视机交互的方法,所述方法包括:
在电子设备处:
从用户接收语音输入;
基于所述语音输入来确定媒体内容;
显示具有第一尺寸的第一用户界面,其中所述第一用户界面包括至所述媒体内容的一个或多个可选择链接;
接收对所述一个或多个可选择链接中的一个可选择链接的选择;以及
响应于所述选择,显示具有比所述第一尺寸更大的第二尺寸的第二用户界面,其中所述第二用户界面包括与所述选择相关联的所述媒体内容。
2. 根据权利要求1所述的方法,其中所述第一用户界面响应于所述选择而扩展到所述第二用户界面中。
3. 根据权利要求1所述的方法,其中所述第一用户界面重叠在播放媒体内容上。
4. 根据权利要求1所述的方法,其中所述第二用户界面重叠在播放媒体内容上。
5. 根据权利要求1所述的方法,其中所述语音输入包括查询,并且所述媒体内容包括查询结果。
6. 根据权利要求5所述的方法,其中所述第一用户界面包括除至所述媒体内容的一个或多个可选择链接之外的至所述查询结果的链接。
7. 根据权利要求5所述的方法,其中所述查询包括关于天气的查询,并且所述第一用户界面包括至与关于所述天气的所述查询相关联的媒体内容的链接。
8. 根据权利要求7所述的方法,其中所述查询包括位置,并且至与关于所述天气的所述查询相关联的所述媒体内容的所述链接包括至与所述位置处的天气相关联的媒体内容的一部分的链接。
9. 根据权利要求1所述的方法,还包括:
响应于所述选择,播放与所述选择相关联的所述媒体内容。
10. 根据权利要求1所述的方法,其中所述媒体内容包括电影。
11. 根据权利要求1所述的方法,其中所述媒体内容包括电视机节目。
12. 根据权利要求1所述的方法,其中所述媒体内容包括体育赛事。
13. 根据权利要求1所述的方法,其中所述第二用户界面包括对与所述选择相关联的所述媒体内容的描述。
14. 根据权利要求1所述的方法,其中所述第一用户界面包括至购买媒体内容的链接。
15. 根据权利要求1所述的方法,还包括:
从所述用户接收附加语音输入,其中所述附加语音输入包括与所显示的内容相关联的查询;
基于与所显示的内容相关联的来确定对与所显示的内容相关联的所述查询的响应;以及
响应于接收到所述附加语音输入,显示第三用户界面,其中所述第三用户界面包括对与所显示的内容相关联的所述查询的所确定的响应。
16. 根据权利要求1所述的方法,还包括:
接收用于发起对语音输入的接收的指示;以及

响应于接收到所述指示,显示准备就绪确认。

17. 根据权利要求1所述的方法,还包括:

响应于接收到所述语音输入,显示侦听确认。

18. 根据权利要求1所述的方法,还包括:

检测所述语音输入的结束;以及

响应于检测到所述语音输入的所述结束,显示处理确认。

19. 根据权利要求1所述的方法,还包括:

显示对所述语音输入的转录。

20. 根据权利要求1所述的方法,其中所述电子设备包括电视机。

21. 根据权利要求1所述的方法,其中所述电子设备包括电视机顶盒。

22. 根据权利要求1所述的设备,其中所述电子设备包括遥控器。

23. 根据权利要求1所述的设备,其中所述电子设备包括移动电话。

24. 根据权利要求1所述的方法,其中所述第一用户界面中的所述一个或多个可选择链接包括与所述媒体内容相关联的移动图像。

25. 根据权利要求24所述的方法,其中与所述媒体内容相关联的所述移动图像包括对所述媒体内容的实况馈送。

26. 根据权利要求1所述的方法,其中所述第一用户界面中的所述一个或多个可选择链接包括与所述媒体内容相关联的静态图像。

27. 根据权利要求1所述的方法,还包括:

确定当前所显示的内容是否包括移动图像或控制菜单;

响应于确定所述当前所显示的内容包括移动图像,选择小尺寸作为所述第一用户界面的所述第一尺寸;以及

响应于确定所述当前所显示的内容包括控制菜单,选择大于所述小尺寸的大尺寸作为所述第一用户界面的所述第一尺寸。

28. 根据权利要求1所述的方法,还包括:

基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示;以及

显示包括所确定的替代媒体内容的通知。

29. 一种非暂态计算机可读存储介质,所述非暂态计算机可读存储介质包含用于执行根据权利要求1-28中任一项所述的方法的计算机可执行指令。

30. 一种系统,包括:

根据权利要求29所述的非暂态计算机可读存储介质;和

能够执行所述计算机可执行指令的处理器。

31. 一种用于使用虚拟助理来控制电视机交互的系统,所述系统包括:

用于从用户接收语音输入的装置;

用于基于所述语音输入来确定媒体内容的装置;

用于显示具有第一尺寸的第一用户界面的装置,其中所述第一用户界面包括至所述媒体内容的一个或多个可选择链接;

用于接收对所述一个或多个可选择链接中的一个可选择链接的选择的装置;和

用于响应于所述选择而显示具有比所述第一尺寸更大的第二尺寸的第二用户界面的装置,其中所述第二用户界面包括与所述选择相关联的所述媒体内容。

32. 一种用于使用虚拟助理来控制电视机交互的方法,所述方法包括:

在电子设备处:

在具有第一显示器的第一设备处从用户接收语音输入;

基于在所述第一显示器上所显示的内容来确定所述语音输入的用户意图;

基于所述用户意图来确定媒体内容;以及

在与第二显示器相关联的第二设备上播放所述媒体内容。

33. 根据权利要求32所述的方法,其中所述第一设备包括遥控器。

34. 根据权利要求32所述的方法,其中所述第一设备包括移动电话。

35. 根据权利要求32所述的方法,其中所述第一设备包括平板电脑。

36. 根据权利要求32所述的方法,其中所述第二设备包括电视机顶盒。

37. 根据权利要求32所述的方法,其中所述第二显示器包括电视机。

38. 根据权利要求32所述的方法,其中在所述第一显示器上所显示的所述内容包括应用界面。

39. 根据权利要求38所述的方法,其中所述语音输入包括用于显示与所述应用界面相关联的媒体的请求。

40. 根据权利要求39所述的方法,其中所述媒体内容包括与所述应用界面相关联的所述媒体。

41. 根据权利要求40所述的方法,其中所述应用界面包括相册,并且所述媒体包括所述相册中的一个或多个照片。

42. 根据权利要求40所述的方法,其中所述应用界面包括一个或多个视频的列表,并且所述媒体包括所述一个或多个视频中的一个视频。

43. 根据权利要求40所述的方法,其中所述应用界面包括电视机节目列表,并且所述媒体包括所述电视机节目列表中的电视机节目。

44. 根据权利要求32所述的方法,还包括:

确定所述第一设备是否被授权;

其中响应于确定所述第一设备被授权而在所述第二设备上播放所述媒体内容。

45. 根据权利要求32所述的方法,还包括:

基于所述语音输入来识别所述用户;以及

基于与所识别的用户相关联的数据来确定所述语音输入的所述用户意图。

46. 根据权利要求32所述的方法,还包括:

基于所述语音输入来确定所述用户是否是被授权;

其中响应于确定所述用户为授权用户而在所述第二设备上播放所述媒体内容。

47. 根据权利要求46所述的方法,其中确定所述用户是否被授权包括使用语音识别来分析所述语音输入。

48. 根据权利要求32所述的方法,还包括:

响应于确定所述用户意图包括对信息的请求,在所述第一设备的所述第一显示器上显示与所述媒体内容相关联的信息;以及

响应于确定所述用户意图包括用于播放所述媒体内容的请求，在所述第二设备上播放所述媒体内容。

49. 根据权利要求32所述的方法，其中所述语音输入包括用于在所述第二设备上播放内容的请求；并且

其中响应于用于在所述第二设备上播放内容的所述请求而在所述第二设备上播放所述媒体内容。

50. 根据权利要求32所述的方法，还包括：

基于媒体格式、用户偏好或默认设置来确定是否应当在所述第一显示器或所述第二显示器上显示所确定的媒体内容；

其中响应于确定应当在所述第二显示器上显示所确定的媒体内容，在所述第二显示器上显示所述媒体内容；

其中响应于确定应当在所述第一显示器上显示所确定的媒体内容，在所述第一显示器上显示所述媒体内容。

51. 根据权利要求32所述的方法，还包括：

确定包括所述第二设备和第三设备的两个或更多个设备中的每个设备的邻近性；

其中基于相对于所述第三设备的所述邻近性的所述第二设备的所述邻近性，在与所述第二显示器相关联的所述第二设备上播放所述媒体内容。

52. 根据权利要求51所述的方法，其中确定所述两个或更多个设备中的每个设备的所述邻近性包括基于蓝牙LE来确定所述邻近性。

53. 根据权利要求32所述的方法，还包括：

显示包括与所述第二显示器相关联的所述第二设备的显示设备的列表；以及

接收对显示设备的所述列表中的所述第二设备的选择；

其中响应于接收到对所述第二设备的所述选择而在所述第二显示器上显示所述媒体内容。

54. 根据权利要求32所述的方法，还包括：

确定耳机是否附接到所述第一设备；

响应于确定所述耳机附接到所述第一设备，在所述第一显示器上显示所述媒体内容；以及

响应于确定所述耳机未附接到所述第一设备，在所述第二显示器上显示所述媒体内容。

55. 根据权利要求32所述的方法，还包括：

基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示；以及

显示包括所确定的替代媒体内容的通知。

56. 一种非暂态计算机可读存储介质，所述非暂态计算机可读存储介质包含用于执行根据权利要求32-55中任一项所述的方法的计算机可执行指令。

57. 一种系统，包括：

根据权利要求56所述的非暂态计算机可读存储介质；和

能够执行所述计算机可执行指令的处理器。

58.一种用于使用虚拟助理来控制电视机交互的系统,所述系统包括:

用于在具有第一显示器的第一设备处从用户接收语音输入的装置;

用于基于在所述第一显示器上所显示的内容来确定所述语音输入的用户意图的装置;

用于基于所述用户意图来确定媒体内容的装置;和

用于在与第二显示器相关联的第二设备上播放所述媒体内容的装置。

59.一种用于使用虚拟助理来控制电视机交互的方法,所述方法包括:

在电子设备处:

从用户接收语音输入,其中所述语音输入包括与在电视机显示器上所显示的内容相关联的查询;

基于在所述电视机显示器上所显示的所述内容和媒体内容的观看历史中的一者或更多者来确定所述查询的用户意图;以及

基于所确定的用户意图来显示查询结果。

60.根据权利要求59所述的方法,其中所述语音输入在遥控器处被接收。

61.根据权利要求59所述的方法,其中所述语音输入在移动电话处被接收。

62.根据权利要求59所述的方法,其中所述查询结果在所述电视机显示器上被显示。

63.根据权利要求59所述的方法,其中在所述电视机显示器上所显示的所述内容包括电影。

64.根据权利要求59所述的方法,其中在所述电视机显示器上所显示的所述内容包括电视机节目。

65.根据权利要求59所述的方法,其中在所述电视机显示器上所显示的所述内容包括体育赛事。

66.根据权利要求59所述的方法,其中所述查询包括对关于与在所述电视机显示器上所显示的所述内容相关联的人的信息的请求,并且所述查询结果包括关于所述人的信息。

67.根据权利要求66所述的方法,其中所述查询结果包括与所述人相关联的媒体内容。

68.根据权利要求67所述的方法,其中所述媒体内容包括与所述人相关联的电影、电视机节目或体育赛事中的一者或更多者。

69.根据权利要求59所述的方法,其中所述查询包括对关于在所述电视机显示器上所显示的所述内容中的角色的信息的请求,并且所述查询结果包括关于所述角色的信息或关于饰演所述角色的演员的信息。

70.根据权利要求69所述的方法,其中所述查询结果包括与饰演所述角色的所述演员相关联的媒体内容。

71.根据权利要求70所述的方法,其中所述媒体内容包括与饰演所述角色的所述演员相关联的电影、电视机节目或体育赛事中的一者或更多者。

72.根据权利要求59所述的方法,还包括:

基于与在所述电视机显示器上所显示的所述内容相关联的元数据或媒体内容的所述观看历史来确定所述查询结果。

73.根据权利要求72所述的方法,其中所述元数据包括以下各项中的一者或更多者:标题、描述、角色列表、演员列表、队员列表、流派、或与在所述电视机显示器上所显示的所述内容相关联的显示调度或媒体内容的所述观看历史。

74. 根据权利要求59所述的方法,其中在所述电视机显示器上所显示的所述内容包括媒体内容列表,并且所述查询包括用于显示所述列表中的项目中的一个项目的请求。

75. 根据权利要求74所述的方法,其中在所述电视机显示器上所显示的所述内容进一步包括所述媒体内容列表中的具有焦点的项目;并且

其中确定所述查询的所述用户意图包括识别具有焦点的所述项目。

76. 根据权利要求59所述的方法,还包括:

基于在所述电视机显示器上最近显示的菜单或搜索内容来确定所述查询的所述用户意图。

77. 根据权利要求76所述的方法,其中在所述电视机显示器上所显示的所述内容包括一页列出的媒体,并且所述最近显示的菜单或搜索内容包括前一页列出的媒体。

78. 根据权利要求59所述的方法,其中在所述电视机显示器上所显示的所述内容包括一种或多种类别的媒体,并且所述一种或多种类别的媒体中的一种类别的媒体具有焦点。

79. 根据权利要求78所述的方法,还包括:

基于所述一种或多种类别的媒体中的所述一种类别的媒体具有焦点来确定所述查询的所述用户意图。

80. 根据权利要求78所述的方法,其中所述媒体的类别包括电影、电视机节目和音乐。

81. 根据权利要求59所述的方法,还包括:

基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示;以及

显示包括所确定的替代媒体内容的通知。

82. 一种非暂态计算机可读存储介质,所述非暂态计算机可读存储介质包含用于执行根据权利要求59-81中任一项所述的方法的计算机可执行指令。

83. 一种系统,包括:

根据权利要求82所述的非暂态计算机可读存储介质;和

能够执行所述计算机可执行指令的处理器。

84. 一种用于使用虚拟助理来控制电视机交互的系统,所述系统包括:

用于从用户接收语音输入的装置,其中所述语音输入包括与在电视机显示器上所显示的内容相关联的查询;

用于基于在所述电视机显示器上所显示的所述内容和媒体内容的观看历史中的一者或者多者来确定所述查询的用户意图的装置;和

用于基于所确定的用户意图来显示查询结果的装置。

85. 一种用于建议虚拟助理交互以用于控制媒体内容的方法,所述方法包括:

在电子设备处:

在显示器上显示媒体内容;

从用户接收输入;

基于所述媒体内容和媒体内容的观看历史中的一者或者多者来确定一个或多个虚拟助理查询;以及

在所述显示器上显示所述一个或多个虚拟助理查询。

86. 根据权利要求85所述的方法,其中所述输入在遥控器上从所述用户接收。

87. 根据权利要求85所述的方法,其中所述输入在移动电话上从所述用户接收。
88. 根据权利要求85所述的方法,其中所述一个或多个虚拟助理查询重叠在移动图像上。
89. 根据权利要求85所述的方法,其中所述输入包括对按钮的双击。
90. 根据权利要求85所述的方法,其中所述媒体内容包括电影。
91. 根据权利要求85所述的方法,其中所述媒体内容包括电视机节目。
92. 根据权利要求85所述的方法,其中所述媒体内容包括体育赛事。
93. 根据权利要求85所述的方法,其中所述一个或多个虚拟助理查询包括关于在所述媒体内容中出现的人的查询。
94. 根据权利要求85所述的方法,其中所述一个或多个虚拟助理查询包括关于在所述媒体内容中出现的角色的查询。
95. 根据权利要求85所述的方法,其中所述一个或多个虚拟助理查询包括对与在所述媒体内容中出现的人相关联的媒体内容的查询。
96. 根据权利要求85所述的方法,其中所述媒体内容或媒体内容的所述观看历史包括电视机节目的剧集,并且所述一个或多个虚拟助理查询包括关于所述电视机节目的另一剧集的查询。
97. 根据权利要求85所述的方法,其中所述媒体内容或媒体内容的所述观看历史包括电视机节目的剧集,并且所述一个或多个虚拟助理查询包括用于设置提醒以观看或录制所述媒体内容的后续剧集的请求。
98. 根据权利要求85所述的方法,其中所述一个或多个虚拟助理查询包括对所述媒体内容的描述性详情的查询。
99. 根据权利要求98所述的方法,其中所述描述性详情包括以下各项中的一者或多者:节目标题、角色列表、演员列表、剧集描述、团队名单、团队评级或节目概要。
100. 根据权利要求85所述的方法,还包括:
接收对所述一个或多个虚拟助理查询中的一个虚拟助理查询的选择;以及
显示所述一个或多个虚拟助理查询中的所选择的一个虚拟助理查询的结果。
101. 根据权利要求85所述的方法,其中确定所述一个或多个虚拟助理查询包括基于以下各项中的一者或多者来确定所述一个或多个虚拟助理查询:查询历史、用户偏好或查询流行性。
102. 根据权利要求85所述的方法,其中确定所述一个或多个虚拟助理查询包括基于所述用户可用以用于进行观看的媒体内容来确定所述一个或多个虚拟助理查询。
103. 根据权利要求85所述的方法,其中确定所述一个或多个虚拟助理查询包括基于所接收的通知来确定所述一个或多个虚拟助理查询。
104. 根据权利要求85所述的方法,其中确定所述一个或多个虚拟助理查询包括基于活动应用程序来确定所述一个或多个虚拟助理查询。
105. 根据权利要求85所述的方法,还包括:
基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示;以及
显示包括所确定的替代媒体内容的通知。

106. 一种非暂态计算机可读存储介质,所述非暂态计算机可读存储介质包含用于执行根据权利要求85-105中任一项所述的方法的计算机可执行指令。

107. 一种系统,包括:

根据权利要求106所述的非暂态计算机可读存储介质;和
能够执行所述计算机可执行指令的处理器。

108. 一种用于建议虚拟助理交互以用于控制媒体内容的系统,所述系统包括:

用于在显示器上显示媒体内容的装置;
用于从用户接收输入的装置;
用于基于所述媒体内容和媒体内容的观看历史中的一者或者来确定一个或多个虚拟助理查询的装置;和
用于在所述显示器上显示所述一个或多个虚拟助理查询的装置。

用于电视机用户交互的智能自动化助理

[0001] 相关申请的交叉引用

[0002] 本申请要求2014年6月30日提交的名称为“INTELLIGENT AUTOMATED ASSISTANT FOR TV USER INTERACTIONS”(用于电视机用户交互的智能自动化助理)的美国临时申请No.62/019,312和2014年9月26日提交的名称为“INTELLIGENT AUTOMATED ASSISTANT FOR TV USER INTERACTIONS”(用于电视机用户交互的智能自动化助理)的美国非临时申请No.14/498,503的优先权,在此出于所有目的通过引用将其全文并入本文。

[0003] 本申请还与以下共同未决的临时申请相关:2014年6月30日提交的名称为“REAL-TIME DIGITAL ASSISTANT KNOWLEDGE UPDATES”(实时数字助理知识更新)的美国专利申请No.62/019,292(代理文献号No.106843097900P22498USP1)),在此通过引用将其全文并入本文。

技术领域

[0004] 本发明整体涉及控制电视机用户的交互,更具体地讲,涉及为虚拟助理处理语音以控制电视机用户的交互。

背景技术

[0005] 智能自动化助理(或虚拟助理)提供用户和电子设备之间的直观界面。这些助理能够允许用户利用口语和/或文本形式的自然语言与设备或系统交互。例如,用户能够通过向与电子设备相关联的虚拟助理提供自然语言形式的口语用户输入来访问电子设备的服务。虚拟助理能够对口语用户输入执行自然语言处理以推断用户的意图并将用户意图操作化成任务。然后可以通过执行电子设备的一个或多个功能来执行任务,并且在一些示例中,可以通过自然语言的形式向用户返回相关输出。

[0006] 尽管移动电话(例如,智能电话)、平板电脑等已经受益于虚拟助理控制,但很多其他用户设备缺乏此类便利的控制机制。例如,用户与媒体控制设备(例如,电视机、电视机顶盒、电缆盒、游戏设备、流媒体设备、数字视频录像机等)的交互可能很复杂且难以学习。此外,随着可以通过此类设备获得的媒体源的增多(例如,空中传播电视机、订阅电视机服务、流视频服务、按需有线视频服务、基于Web的视频服务等),一些用户要找到期望的媒体内容来消费可能会很繁琐,甚至无法进行。结果,很多媒体控制设备可能提供欠佳的用户体验,这会使很多用户感到沮丧。

发明内容

[0007] 本发明公开了用于使用虚拟助理来控制电视机交互的系统和过程。在一个实施例中,可以从用户接收语音输入。可基于语音输入来确定媒体内容。可以显示具有第一尺寸的第一用户界面,该第一用户界面可包括至媒体内容的可选择链接。可以接收对可选择链接中的一个可选择链接的选择。响应于该选择,可以显示具有比第一尺寸更大的第二尺寸的第二用户界面,第二用户界面可包括与该选择相关联的媒体内容。

[0008] 在另一个实施例中,可以在具有第一显示器的第一设备处从用户接收语音输入。可基于在第一显示器上所显示的内容来确定语音输入的用户意图。可基于用户意图来确定媒体内容。可以在与第二显示器相关联的第二设备上播放媒体内容。

[0009] 在另一个实施例中,可以从用户接收语音输入,该语音输入可包括与在电视机显示器上所显示的内容相关联的查询。可基于在电视机显示器上所显示的内容和/或媒体内容的查看历史来确定查询的用户意图。可基于确定的用户意图来显示查询结果。

[0010] 在另一个实施例中,可以在显示器上显示媒体内容。可以从用户接收输入。可基于媒体内容和/或媒体内容的查看历史来确定虚拟助理查询。可以在显示器上显示虚拟助理查询。

附图说明

- [0011] 图1示出了用于使用虚拟助理来控制电视机用户交互的示例性系统。
- [0012] 图2示出了根据各种实施例的示例性用户设备的框图。
- [0013] 图3示出了用于控制电视机用户交互的系统中的示例性媒体控制设备的框图。
- [0014] 图4A-图4E示出了视频内容上方的示例性语音输入界面。
- [0015] 图5示出了视频内容上方的示例性媒体内容界面。
- [0016] 图6A-图6B示出了视频内容上方的示例性媒体详情界面。
- [0017] 图7A-图7B示出了示例性媒体转变界面。
- [0018] 图8A-图8B示出了菜单内容上方的示例性语音输入界面。
- [0019] 图9示出了菜单内容上方的示例性虚拟助理结果界面。
- [0020] 图10示出了用于使用虚拟助理来控制电视机交互和使用不同界面来显示关联信息的示例性过程。
- [0021] 图11示出了移动用户设备上的示例性电视机媒体内容。
- [0022] 图12示出了使用虚拟助理的示例性电视机控制。
- [0023] 图13示出了移动用户设备上的示例性画面和视频内容。
- [0024] 图14示出了使用虚拟助理的示例性媒体显示控制。
- [0025] 图15示出了在移动用户设备和媒体显示设备上具有结果的示例性虚拟助理交互。
- [0026] 图16示出了在媒体显示设备和移动用户设备上具有媒体结果的示例性虚拟助理交互。
- [0027] 图17示出了基于接近性的示例性媒体设备控制。
- [0028] 图18示出了用于使用虚拟助理和多个用户设备来控制电视机交互的示例性过程。
- [0029] 图19示出了具有关于背景视频内容的虚拟助理查询的示例性语音输入界面。
- [0030] 图20示出了视频内容上方的示例性信息虚拟助理。
- [0031] 图21示出了具有用于与背景视频内容相关联的媒体内容的虚拟助理查询的示例性语音输入界面。
- [0032] 图22示出了具有可选择媒体内容的示例性虚拟助理响应界面。
- [0033] 图23A-图23B示出了程序菜单的示例性页面。
- [0034] 图24示出了分成类别的示例性媒体菜单。
- [0035] 图25示出了用于利用显示器上所显示的媒体内容控制电视机交互和媒体内容的

查看历史的示例性过程。

- [0036] 图26示出了基于背景视频内容的具有虚拟助理查询建议的示例性界面。
- [0037] 图27示出了用于确认所建议查询选择的示例性界面。
- [0038] 图28A-图28B示出了基于所选择的查询的示例性虚拟助理应答界面。
- [0039] 图29示出了媒体内容通知和具有基于该通知的虚拟助理查询建议的示例性界面。
- [0040] 图30示出了具有示例性图画和可在媒体控制设备上播放的视频内容的移动用户设备。
- [0041] 图31示出了具有基于可播放用户设备内容且基于独立显示器上所显示的视频内容的虚拟助理查询建议的示例性移动用户设备界面。
- [0042] 图32示出了具有基于来自独立用户设备的可播放内容的虚拟助理查询建议的示例性界面。
- [0043] 图33示出了用于建议虚拟助理交用于互以控制媒体内容的示例性过程。
- [0044] 图34示出了电子设备的功能框图,该电子设备被配置为使用虚拟助理来控制电视机交互并利用根据各种实施例的不同界面显示关联信息。
- [0045] 图35示出了电子设备的功能框图,该电子设备被配置为使用虚拟助理和根据各种实施例的多个用户设备来控制电视机交互。
- [0046] 图36示出了电子设备的功能框图,该电子设备被配置为利用显示器上所显示的媒体内容以及根据各种实施例的媒体内容的查看历史来控制电视机交互。
- [0047] 图37示出了电子设备的功能框图,该电子设备被配置为建议虚拟助理交互,用于根据各种实施例控制媒体内容。

具体实施方式

[0048] 在以下对实施例的描述中将引用附图,在附图中以例示的方式示出了可被实施的特定示例。应当理解,在不脱离各种实施例的范围的情况下,可使用其他实施例并且可作出结构性变更。

[0049] 本发明涉及用于使用虚拟助理来控制电视机用户交互的系统和过程。在一个实施例中,可以使用虚拟助理与媒体控制设备交互,例如控制在电视机显示器上所显示的内容的电视机顶盒。可以使用移动用户设备或具有麦克风的遥控器接收针对虚拟助理的语音输入。可以从语音输入确定用户的意图,虚拟助理可以根据用户的意图执行任务,包括使得在连接的电视机上回放媒体以及控制电视机顶盒或类似设备的任何其他功能(例如,管理视频录制、搜索媒体内容、在菜单中导航等)。

[0050] 可以在连接的电视机或其他显示器上显示虚拟助理交互。在一个实施例中,可基于从用户接收的语音输入来确定媒体内容。可以显示具有第一小尺寸的第一用户界面,包括至所确定的媒体内容的可选择链接。在接收对媒体链接的选择之后,可以显示具有第二更大尺寸的第二用户界面,包括与该选择相关联的媒体内容。在其他实施例中,用于传输虚拟助理交互的界面可以扩展或收缩以占据最小量的空间,同时传输期望的信息。

[0051] 在一些实施例中,可以使用与多个显示器相关联的多个设备从语音输入确定用户意图,以及通过不同方式向用户传输信息。例如,可以在具有第一显示器的第一设备处从用户接收语音输入。可基于在第一显示器上所显示的内容从语音输入确定用户意图。可基于

用户意图来确定媒体内容，并可以在与第二显示器相关联的第二设备上播放媒体内容。

[0052] 还可以使用电视机显示内容作为用于从语音输入确定用户意图的语境输入。例如，可以从用户接收语音输入，包括与在电视机显示器上所显示的内容相关联的查询。可基于在电视机显示器上所显示的内容以及电视机显示器上的媒体内容的查看历史来确定查询的用户意图（例如，基于播放电视机节目中的演员消除查询的歧义）。然后可基于确定的用户意图来显示查询结果。

[0053] 在一些实施例中，可以向用户提供虚拟助理查询建议（例如，通知用户有可用的命令，建议感兴趣内容等）。例如，可以在显示器上显示媒体内容，并可以从用户接收请求虚拟助理查询建议的输入。可基于显示器上所显示的媒体内容和显示器上所显示的媒体内容的查看历史来确定虚拟助理查询建议（例如，建议与播放电视机节目相关的查询）。然后可以在显示器上显示建议的虚拟助理查询。

[0054] 根据本文论述的各种实施例使用虚拟助理来控制电视机用户交互能够提供有效率且享受性的用户体验。使用能够接收自然语言查询或命令的虚拟助理，用户与媒体控制设备的交互可以是直观且简单的。可以根据需要向用户建议可用功能，包括基于播放内容的有意义查询建议，这样能够辅助用户学习控制能力。此外，可以利用直观的口语命令使得可用媒体容易访问。然而，应当理解，根据本文论述的各种实施例，仍然能够实现实现很多其他优点。

[0055] 图1示出了用于使用虚拟助理来控制电视机用户交互的示例性系统100。应当理解，如本文所述那样控制电视机用户交互仅仅是控制一种显示技术上的媒体的一个示例且用于参考，可以将本文所述的概念一般地用于控制任何媒体内容交互，包括各种设备和关联显示器上的任何内容（例如，监视器、膝上型显示器、台式计算机显示器、移动用户设备显示器、投影仪显示器等）。因此，术语“电视机”可指与各种设备的任一种相关联的任何类型的显示器。此外，术语“虚拟助理”、“数字助理”、“智能自动化助理”或“自动数字助理”可指解译口头和/或文本形式的自然语言输入以推断用户意图，并基于推断出的用户意图来执行动作的任何信息处理系统。例如，为了按推断出的用户意图执行动作，系统可执行以下内容中的一个或多个：通过利用专为实现所推断出的用户意图设计的步骤和参数来识别任务流；将来自推断出的用户意图的具体要求输入到任务流中；通过调用程序、方法、服务、API等来执行任务流；以及生成对用户的听觉（例如，语音）和/或视觉形式的输出响应。

[0056] 虚拟助理能够接受至少部分地为自然语言命令、请求、声明、讲述和/或询问形式的用户请求。通常，用户请求要么寻求通过虚拟助理作出信息性回答，要么寻求通过虚拟助理执行任务（例如，使得显示特定媒体）。针对用户请求的令人满意的响应可包括提供所请求的信息性回答、执行所请求的任务或这两者的组合。例如，用户可向虚拟助理提出诸如“我现在在哪里？”之类的问题。基于用户的当前位置，虚拟助理可能回答“你在中央公园。”用户还可以请求执行任务，例如，“请提醒我今天下午4点给妈妈打电话。”作为响应，虚拟助理能够确认该请求，然后在用户的电子日程中创建适当的提醒项。在执行所请求的任务期间，虚拟助理可有时在很长时间段内在涉及多次信息交换的持续对话中与用户进行交互。存在与虚拟助理进行交互以请求信息或执行各种任务的许多其他方法。除提供言语响应并采取经编程的动作之外，虚拟助理还可提供其他视觉或音频形式（例如，作为文本、警报、音乐、视频、动画等）的响应。此外，如本文所述，示例性虚拟助理能够控制媒体内容的回放（例

如,在电视机上播放视频)并使得在显示器上显示信息。

[0057] 虚拟助理的示例在申请人2011年1月10日提交的名称为“Intelligent Automated Assistant”(智能自动化助理)的美国实用新型专利申请No.12/987,982中有所描述,其整个公开内容以引用方式并入本文。

[0058] 如图1中所示,在一些实施例中,虚拟助理可根据客户端-服务器模型来实施。虚拟助理可包括在用户设备102上执行的客户端侧部分和在服务器系统110上执行的服务器侧部分。客户端侧部分也可以结合遥控器106在电视机顶盒104上执行。用户设备102可包括任何电子设备,例如移动电话(例如,智能电话)、平板电脑、便携式媒体播放器、台式计算机、膝上型计算机、PDA、可穿戴电子设备(例如,数字眼镜、腕带、腕表、胸针、臂带等)等。电视机顶盒104可包括任何媒体控制设备,例如,电缆盒、卫星盒、视频播放器、视频流传输设备、数字视频录像机、游戏系统、DVD播放机、蓝光光盘TM播放器、此类设备的组合等。电视机顶盒104可以经由有线或无线连接而连接到显示器112和扬声器111。显示器112(具有或没有扬声器111)可以是任何类型的显示器,例如电视机显示器、监视器、投影仪等。在一些实施例中,电视机顶盒104可以连接到音频系统(例如,音频接收机),扬声器111可以独立于显示器112。在其他实施例中,显示器112、扬声器111和电视机顶盒104可以一起并入单个设备中,例如具有高级处理和网络连接能力的智能电视机。在此类实施例中,电视机顶盒104的功能可以被执行为组合设备上的应用。

[0059] 在一些实施例中,电视机顶盒104可以充当用于多种类型和来源的媒体内容的媒体控制中心。例如,电视机顶盒104能够方便用户访问实况电视机(例如,空中传播、卫星或有线电视机)。这样一来,电视机顶盒104可包括电缆调谐器、卫星调谐器等。在一些实施例中,电视机顶盒104还可以记录电视机节目,供稍晚其他时间观看。在其他实施例中,电视机顶盒104可以提供对一个或多个流媒体服务的访问,例如电缆传输的按需电视机节目、视频和音乐,以及互联网传输的电视机节目、视频和音乐(例如,来自各种免费、付费和基于订阅的流媒体服务)。在其他示例中,电视机顶盒104可以方便从任何其他源回放或显示媒体内容,例如从移动用户设备显示照片,从耦接的存储设备播放视频、从耦接的音乐播放器播放音乐等。电视机顶盒104还可以根据需要包括本文论述的媒体控制特征的各种其他组合。

[0060] 用户设备102和电视机顶盒104能够通过一个或多个网络108与服务器系统110通信,网络可包括互联网、内联网或任何其他有线或无线公共或私有网络。此外,用户设备102能够通过网络108或直接通过任何其他有线或无线通信机制(例如,蓝牙、Wi-Fi、射频、红外传输等)与电视机顶盒104通信。如图所示,遥控器106能够使用任何类型的通信,例如有线连接或任何类型的无线通信(例如,蓝牙、Wi-Fi、射频、红外传输等),包括经由网络108,与电视机顶盒104通信。在一些示例中,用户能够通过用户设备102、遥控器106或集成于电视机顶盒104之内的接口元件(例如,按钮、麦克风、相机、操纵杆等)与电视机顶盒104交互。例如,可以在用户设备102和/或遥控器106处接收包括用于虚拟助理的媒体相关查询或命令的语音输入,并可以使用语音输入使得媒体相关的任务在电视机顶盒104上执行。同样,可以在用户设备102和/或遥控器106处(以及从未示出的其他设备)接收用于控制电视机顶盒104上的媒体的触觉命令。因此,可以通过各种方式控制电视机顶盒104的各种功能,为用户赋予用于控制来自多个设备的媒体内容的多个选项。

[0061] 在用户设备102和/或具有遥控器106的电视机顶盒104上执行的示例性虚拟助理

的客户端侧部分能够提供客户端侧功能,例如面向用户的输入和输出处理以及与服务器系统110的通信。服务器系统110可以为驻留在相应用户设备102或相应电视机顶盒104上的任意数量的客户端提供服务器侧功能。

[0062] 服务器系统110可包括一个或多个可包括面向客户端的I/O接口122的虚拟助理服务器114、一个或多个处理模块118、数据与模型存储装置120,以及到外部服务的I/O接口116。面向客户端的I/O接口122可促进针对虚拟助理服务器114的面向客户端的输入和输出处理。一个或多个处理模块118可利用数据与模型存储装置120基于自然语言输入来确定用户的意图,并基于推断出的用户意图来进行任务执行。在一些示例中,虚拟助理服务器114能够通过网络108与外部服务124通信,外部服务124例如是电话服务、日历服务、信息服务、即时消息服务、导航服务、电视机节目服务、流媒体服务等,以完成任务或获取信息。到外部服务的I/O接口116可促进此类通信。

[0063] 服务器系统110可在计算机的一个或多个独立式数据处理设备或分布式网络上实施。在一些示例中,服务器系统110可采用第三方服务提供商(例如,第三方云服务提供商)的各种虚拟设备和/或服务来提供服务器系统110的底层计算资源和/或基础结构资源。

[0064] 尽管图1中将虚拟助理的功能显示为包括客户端侧部分和服务器侧部分两者,但在一些示例中,可以将助理的功能(或一般的语音识别和媒体控制)实现为安装于用户设备、电视机顶盒、智能电视机等上的独立应用。此外,虚拟助理的客户端部分与服务器部分之间的功能划分在不同的示例中可变化。例如,在一些示例中,用户设备102或电视机顶盒104上执行的客户端可以是瘦客户端,其仅提供面向用户的输入和输出处理功能,并将虚拟助理的所有其他功能委托给后端服务器。

[0065] 图2示出了根据各种实施例的示例性用户设备102的框图。如图所示,用户设备102可包括存储器接口202、一个或多个处理器204,以及外围设备接口206。用户设备102中的各种部件可由一条或多条通信总线或信号线耦接在一起。用户设备102可进一步包括各种传感器、子系统以及耦接至外围设备接口206的外围设备。传感器、子系统以及外围设备采集信息并且/或者促进用户设备102的各种功能。

[0066] 例如,用户设备102可包括运动传感器210、光传感器212以及耦接至外围设备接口206的接近传感器214以促进取向、照明和接近感测功能。一个或多个其他传感器216,诸如定位系统(例如,GPS接收器)、温度传感器、生物计量传感器、陀螺仪、指南针、加速度计等,也可连接至外围设备接口206,以促进相关功能。

[0067] 在一些示例中,相机子系统220和光学传感器222可用于促进相机功能,诸如拍摄照片和记录视频剪辑。通信功能可通过一个或多个有线和/或无线通信子系统224来促进,该有线和/或无线通信子系统可包括各种通信端口、射频接收器与发射器和/或光学(如红外)接收器与发射器。可将音频子系统226耦接到扬声器228和麦克风230以促进启用语音的功能,诸如语音识别、语音复制、数字记录和电话功能。

[0068] 在一些示例中,用户设备102能够进一步包括耦接到外围设备接口206的I/O子系统240。I/O子系统240可包括触摸屏控制器242和/或其他输入控制器244。触摸屏控制器242可以耦接到触摸屏246。触摸屏246和触摸屏控制器242例如可使用多种触摸灵敏度技术中的任一种来检测接触和移动或它们的间断,该多种触摸灵敏度技术诸如电容技术、电阻技术、红外和表面声波技术、接近传感器阵列等。可将其他输入控制器244耦接到其他输入/控

制设备248,诸如一个或多个按钮、摇臂开关、拇指滚轮、红外端口、USB端口和/或指针设备(诸如触笔)。

[0069] 在一些示例中,用户设备102还可包括耦接到存储器250的存储器接口202。存储器250可包括任意电子、磁性、光学、电磁、红外或半导体系统、装置或设备;便携式计算机磁盘(磁性)、随机存取存储器(RAM)(磁性);只读存储器(ROM)(磁性);可擦除可编程只读存储器(EPROM)(磁性);便携式光盘诸如CD、CD-R、CD-RW、DVD、DVD-R或DVD-RW;或者闪存存储器诸如紧凑型闪存卡、安全数字卡、USB存储器设备、记忆棒等。在一些示例中,存储器250的非暂态计算机可读存储介质可以用于存储指令(例如,用于执行本文描述的各种过程的部分或全部),以供指令执行系统、装置或设备使用或结合其使用,例如基于计算机的系统、包含处理器的系统或可从指令执行系统、装置或设备获取指令并能够执行指令的其他系统。在其他示例中,指令(例如,用于执行本文所述的各种过程的部分或全部)可以存储于服务器系统110的非暂态计算机可读存储介质上,或者可以在存储器250的非暂态计算机可读存储介质和服务器系统110的非暂态计算机可读存储介质之间分配。在本文的语境中,“非暂态计算机可读存储介质”可以是可包括或存储程序以供指令执行系统、装置和设备使用或与其结合的任何介质。

[0070] 在一些示例中,存储器250可存储操作系统252、通信模块254、图形用户界面模块256、传感器处理模块258、电话模块260和应用262。操作系统252可包括用于处理基础系统服务以及用于执行硬件相关任务的指令。通信模块254可促进与一个或多个附加设备、一个或多个计算机和/或一个或多个服务器的通信。图形用户界面指令256可促进图形用户界面处理。传感器处理模块258可促进与传感器相关的处理和功能。电话模块260可促进与电话相关的过程和功能。应用模块262可促进用户应用的各种功能,诸如电子消息传送、网页浏览、媒体处理、导航、成像和/或其他过程和功能。

[0071] 如本文所述,存储器250还可存储客户端侧虚拟助理指令(例如,在虚拟助理客户端模块264中)以及各种用户数据266(例如,特定于用户的词汇数据、偏好数据,和/或其他数据诸如用户的电子通讯录、待办事项列表、购物清单、电视机节目收藏等)以,例如提供虚拟助理的客户端侧功能。用户数据266还可以用于在支持虚拟助理或针对任何其他应用时进行语音识别。

[0072] 在各种示例中,虚拟助理客户端模块264能够通过用户设备102的各种用户界面(例如,I/O子系统240、音频子系统226等)接受声音输入(例如,语音输入)、文本输入、触摸输入和/或手势输入。虚拟助理客户端模块264还能够提供音频(例如,语音输出)、视觉和/或触觉形式的输出。例如,可将输出提供为语音、声音、警报、文本消息、菜单、图形、视频、动画、振动、和/或以上两者或更多者的组合。在操作期间,虚拟助理客户端模块264使用通信子系统224来与虚拟助理服务器通信。

[0073] 在一些示例中,虚拟助理客户端模块264可利用各种传感器、子系统和外围设备来从用户设备102的周围环境采集附加信息以建立与用户、当前用户交互和/或当前用户输入相关联的上语境。此类语境还可包括来自其他设备的信息,例如,来自电视机顶盒104。在一些示例中,虚拟助理客户端模块264可将语境信息或其子集与用户输入一起提供至虚拟助理服务器以帮助推断用户的意图。虚拟助理还可使用上下文信息来确定如何准备输出并将其传送至用户。该语境信息还可以由用户设备102或服务器系统110使用以支持精确的语言

识别。

[0074] 在一些示例中,伴随用户输入的语境信息可包括传感器信息,例如照明、环境噪音、环境温度、周围环境的图像或视频、与另一物体的距离等。语境信息还可包括与用户设备102的物理状态(例如,设备取向、设备位置、设备温度、功率水平、速度、加速度、运动模式、蜂窝信号强度等)或用户设备102的软件状态(例如,运行过程、安装的程序、过去和现在的网络活动、背景服务、错误日志、资源使用等)相关联的信息。语境信息还可包括与连接的设备或与用户相关联的其他设备的状态相关联的信息(例如,电视机顶盒104显示的媒体内容、可用于电视机顶盒104的媒体内容等)。这些类型的语境信息的任一种都可以被提供给虚拟助理服务器114(或用于用户设备102自身上)作为与用户输入相关联的语境信息。

[0075] 在一些示例中,虚拟助理客户端模块264可以响应于来自虚拟助理服务器114的请求,选择性地提供用户设备102上存储的信息(例如,用户数据266)(或者可以在执行语音识别和/或虚拟助理功能时将其用于用户设备102自身上)。虚拟助理客户端模块264还可在虚拟助理服务器114请求时引出来自用户经由自然语言对话或其他用户界面的附加输入。虚拟助理客户端模块264可将该附加输入传递至虚拟助理服务器114,以帮助虚拟助理服务器114进行意图推断和/或满足在用户请求中表达的用户意图。

[0076] 在各种示例中,存储器250可包括附加指令或更少的指令。此外,可在硬件和/或固件中包括在一个或多个信号处理和/或专用集成电路中执行用户设备102的各种功能。

[0077] 图3示出了用于控制电视机用户交互的系统300中的示例性电视机顶盒104的框图。系统300可包括系统100的元件的子集。在一些示例中,系统300可以单独执行特定功能,并能够与系统100的其他元件一起工作以执行其他功能。例如,系统300的元件能够处理特定媒体控制功能而不和服务器系统110交互(例如,回放本地存储的媒体、记录功能、频道调谐等),系统300能够结合服务器系统110和系统100的其他元件处理其他媒体控制功能(例如,回放远程存储的媒体、下载媒体内容、处理特定虚拟助理查询等)。在其他示例中,系统300的元件能够执行更大系统100的功能,包括通过网络访问外部服务124。应当理解,可以通过各种其他方式在本地设备和远程服务器设备之间划分功能。

[0078] 如图3所示,在一个示例中,电视机顶盒104能够包括存储器接口302、一个或多个处理器304和外围设备接口306。电视机顶盒104中的各种部件可由一条或多条通信总线或信号线耦接在一起。电视机顶盒104可进一步包括各种子系统以及耦接至外围设备接口306的外围设备。子系统以及外围设备可采集信息并且/或者促进电视机顶盒104的各种功能。

[0079] 例如,电视机顶盒104可包括通信子系统324。通信功能可通过一个或多个有线和/或无线通信子系统324来促进,该有线和/或无线通信子系统可包括各种通信端口、射频接收器与发射器和/或光学(如红外)接收器与发射器。

[0080] 在一些示例中,电视机顶盒104还可包括耦接到外围设备接口306的I/O子系统340。I/O子系统340可包括音频/视频输出控制器370。音频/视频输出控制器370可以耦接到显示器112和扬声器111或者可以通过其他方式提供音频和视频输出(例如,经由音频/视频端口、无线传输等)。I/O子系统340还可包括远程控制器342。远程控制器342可以通信地耦接到遥控器106(例如,经由有线连接、蓝牙、Wi-Fi等)。遥控器106可包括用于捕获音频输入(例如,来自用户的语音输入)的麦克风372、用于捕获触觉输入的按钮374和用于方便经由远程控制器342与电视机顶盒104通信的收发器376。遥控器106还可包括诸如键盘、操纵杆、

触控板等的其他输入机构。遥控器106还可包括诸如灯、显示器、扬声器等的输出机构。遥控器106处接收的输入(例如,用户语音、按钮按压等)可以经由远程控制器342被传送到电视机顶盒104。I/O子系统340还可包括其他输入控制器344。可将其他输入控制器344耦接到其他输入/控制设备348,诸如一个或多个按钮、摇臂开关、拇指滚轮、红外端口、USB端口和/或指针设备(诸如触笔)。

[0081] 在一些示例中,电视机顶盒104还可包括耦接到存储器350的存储器接口302。存储器350可包括任意电子、磁性、光学、电磁、红外或半导体系统、装置或设备;便携式计算机磁盘(磁性)、随机存取存储器(RAM)(磁性);只读存储器(ROM)(磁性);可擦除可编程只读存储器(EPROM)(磁性);便携式光盘诸如CD、CD-R、CD-RW、DVD、DVD-R或DVD-RW;或者闪存存储器诸如紧凑型闪存卡、安全数字卡、USB存储器设备、记忆棒等。在一些示例中,存储器350的非暂态计算机可读存储介质可以用于存储指令(例如,用于执行本文描述的各种过程的部分或全部),以供指令执行系统、装置或设备使用或结合其使用,例如基于计算机的系统、包含处理器的系统或可从指令执行系统、装置或设备获取指令并能够执行指令的其他系统。在其他示例中,指令(例如,用于执行本文所述的各种过程的部分或全部)可以存储于服务器系统110的非暂态计算机可读存储介质上,或者可以在存储器350的非暂态计算机可读存储介质和服务器系统110的非暂态计算机可读存储介质之间分配。在本文档的语境中,“非暂态计算机可读存储介质”可以是可包括或存储程序以供指令执行系统、装置和设备使用或与其结合的任何介质。

[0082] 在一些示例中,存储器350可存储操作系统352、通信模块354、图形用户界面模块356、设备上媒体模块358、设备外媒体模块360和应用362。操作系统352可包括用于处理基础系统服务以及用于执行硬件相关任务的指令。通信模块354可促进与一个或多个附加设备、一个或多个计算机和/或一个或多个服务器的通信。图形用户界面指令356可促进图形用户界面处理。设备上媒体模块358可以方便电视机顶盒104上本地存储的媒体内容和其他可以在本地获得的媒体内容(例如,有线频道调谐)的存储和回放。设备外媒体模块360能够方便远程存储(例如,在远程服务器上,在用户设备102上等)的媒体内容的流回放或下载。应用模块362可促进用户应用的各种功能,诸如电子消息传送、Web浏览、媒体处理、游戏和/或其他过程和功能。

[0083] 如本文所述,存储器350还可存储客户端侧虚拟助理指令(例如,在虚拟助理客户端模块364中)以及各种用户数据366(例如,特定于用户的词汇数据、偏好数据,和/或其他数据诸如用户的电子通讯录、待办事项列表、购物清单、电视机节目收藏等)以,例如提供虚拟助理的客户端侧功能。用户数据366还可用于在支持虚拟助理或针对任何其他应用时进行语音识别。

[0084] 在各种示例中,虚拟助理客户端模块364能够通过电视机顶盒104的各种用户界面(例如,I/O子系统340等)接受声音输入(例如,语音输入)、文本输入、触摸输入和/或手势输入。虚拟助理客户端模块364还能够提供音频(例如,语音输出)、视觉和/或触觉形式的输出。例如,可将输出提供为语音、声音、警报、文本消息、菜单、图形、视频、动画、振动、和/或以上两者或更多者的组合。在操作期间,虚拟助理客户端模块364使用通信子系统324来与虚拟助理服务器通信。

[0085] 在一些示例中,虚拟助理客户端模块364可利用各种子系统和外围设备来从电视

机顶盒104的周围环境采集附加信息以建立与用户、当前用户交互和/或当前用户输入相关联的上下文。此类语境还可包括来自其他设备,例如,来自用户设备102的信息。在一些示例中,虚拟助理客户端模块364可将上下文信息或其子集与用户输入一起提供至虚拟助理服务器以帮助推断用户的意图。虚拟助理还可使用上下文信息来确定如何准备输出并将其传送至用户。该语境信息还可以由电视机顶盒104或服务器系统110使用以支持精确的语音识别。

[0086] 在一些示例中,伴随用户输入的语境信息可包括传感器信息,例如照明、环境噪音、环境温度、与另一物体的距离等。语境信息还可包括与电视机顶盒104的物理状态(例如,设备位置、设备温度、功率水平等)或电视机顶盒104的软件状态(例如,运行的过程、安装的程序、过去和现在的网络活动、背景服务、错误日志、资源使用等)相关联的信息。语境信息还可包括与连接的设备或与用户相关联的其他设备的状态相关联的信息(例如,在用户设备102上显示的内容、用户设备102上可播放的内容等)。这些类型的语境信息的任一种可以被提供给虚拟助理服务器114(或用于电视机顶盒104自身上)作为与用户输入相关联的语境信息。

[0087] 在一些示例中,虚拟助理客户端模块364可以响应于来自虚拟助理服务器114的请求,选择性地提供电视机顶盒104上存储的信息(例如,用户数据366)(或者可以在执行语音识别和/或虚拟助理功能时将其用于电视机顶盒104自身上)。虚拟助理客户端模块364还可在虚拟助理服务器114请求时引出来自用户经由自然语言对话或其他用户界面的附加输入。虚拟助理客户端模块364可将该附加输入传送至虚拟助理服务器114,以帮助虚拟助理服务器114进行意图推断和/或满足在用户请求中表达的用户意图。

[0088] 在各种示例中,存储器350可包括附加指令或更少的指令。此外,可在硬件和/或固件中包括在一个或多个信号处理和/或专用集成电路中执行电视机顶盒104的各种功能。

[0089] 应当理解,系统100和系统300不限于图1和图3中所示的部件和配置,用户设备102、电视机顶盒104和遥控器106类似地不限于图2和图3中所示的部件和配置。系统100、系统300、用户设备102、电视机顶盒104和遥控器106全部可包括根据各种示例的多种配置中的更少或其他部件。

[0090] 在整个本公开中,提到“系统”可包括系统100、系统300或系统100或系统300的一个或多个元件。例如,本文提到的典型系统可以至少包括从遥控器106和/或用户设备102接收用户输入的电视机顶盒104。

[0091] 图4A到图4E示出了示例性语音输入界面484,其可以显示于显示器(例如,显示器112)上以向用户传送语音输入信息。在一个实施例中,语音输入界面484可以显示于视频480上,其可包括任何移动图像或暂停的视频。例如,视频480可包括实况电视机、播放视频、流媒体电影、录制节目的回放等。语音输入界面484可被配置为占据最少量的空间,以免显著干扰用户观看视频480。

[0092] 在一个实施例中,可以触发虚拟助理以侦听包含命令或查询的语音输入(或开始记录用于后续处理的语音输入,或开始实时处理语音输入)。可以通过各种方式触发侦听,包括指示,例如用户按压遥控器106上的物理按钮、用户按压用户设备102上的物理按钮、用户按压用户设备102上的虚拟按钮、用户发出可以由始终侦听的设备识别的短语(例如,发出“你好,助理”以开始侦听命令)、用户执行可以由传感器检测到的手势(例如,在相机前方

监测)等。在另一个实施例中,用户可以按压并保持遥控器106或用户设备102上的物理按钮以发起侦听。在其他实施例中,用户可以按压和保持遥控器106或用户设备102上的物理按钮,同时讲出查询或命令,并能够在完成时释放按钮。可以类似地接收各种其他指示以发起从用户接收语音输入。

[0093] 响应于接收到侦听语音输入的指示,可以显示语音输入界面484。图4A示出了从显示器112的底部部分向上扩大的通知区域482。可以在接收到侦听语音输入的指示时在通知区域482中显示语音输入界面484,可以动画显示该界面,以从所示的显示器112的观看区域的底边缘向上滑动。图4B示出了向上滑动到查看之后的语音输入界面484。语音输入界面484可被配置为占据显示器112底部最少量的空间,以免显著干扰视频480。响应于接收到侦听语音输入的指示,可以显示准备就绪确认486。准备就绪确认486可包括所示的麦克风符号,或者可包括任何其他图像、图标、动画或符号,以传送系统(例如,系统100的一个或多个元件)准备好从用户捕获语音输入。

[0094] 在用户开始讲话时,可以显示图4C中所示的侦听确认487以确认该系统正在捕获语音输入。在一些实施例中,响应于接收到语音输入(例如,捕获语音),可以显示侦听确认487。在其他实施例中,可以将准备就绪确认486显示预先确定量的时间(例如,500毫秒、1秒、3秒等),之后可以显示侦听确认487。侦听确认487可包括所示的波形符号,或者可包括响应于用户语音而运动(例如,改变频率)的活动波形动画。在其他实施例中,侦听确认487可包括任何其他图像、图标、动画或符号,以传送该系统正在从用户捕获语音输入。

[0095] 在(例如,基于暂停、指示查询结束的语音中断或任何其他端点检测方法)检测到用户已经完成讲话时,可以显示图4D中所示的处理确认488,以确认该系统完成了捕获语音输入,且正在处理语音输入(例如,解释语音输入、确定用户意图和/或执行关联任务)。处理确认488可包括所示的沙漏符号,或者可包括任何其他图像、图标、动画或符号,以传送系统正在处理捕获的语音输入。在另一个实施例中,处理确认488可包括旋转圆圈或在圆圈周围运动的彩色/发光点的动画。

[0096] 在将捕获的语音输入解释为文本(或响应于将语音输入成功转换成文本)之后,可以显示图4E所示的命令接收确认490和/或抄本492,以确认该系统接收并解释了语音输入。抄本492可包括所接收语音输入的抄本(例如,“现在有什么体育赛事?”)。在一些实施例中,可以动画显示抄本492以从显示器112的底部向上滑动,可以在图4E中所示的位置短暂显示抄本(例如,几秒钟),然后可以在从视图中消失之前将其滑动到语音输入界面484顶部(例如,如同文本向上卷动并最终到视图之外那样)。在其他实施例中,可以不显示抄本,可以处理用户的命令或查询,并可以执行关联的任务而不显示抄本(例如,可以立即执行简单的频道改变而不显示用户语音的抄本)。

[0097] 在其他实施例中,可以在用户讲话时实时地执行语音转录。在转录文字时,可以在语音输入界面484中显示它们。例如,可与侦听确认487并排显示文字。在用户完成讲话之后,可以在执行与用户的命令相关联的任务之前简短地显示命令接收确认490。

[0098] 此外,在其他实施例中,命令接收确认490可以传送关于已接收和理解的命令的信息。例如,对于改变到另一频道的简单请求,可以在改变频道时将与频道相关联的徽标或数字简短显示为命令接收确认490(例如,几秒钟)。在另一个实施例中,对于暂停视频的请求(例如,视频480),可以将暂停符号(例如,两个垂直平行条)显示为命令接收确认490。暂停

符号可以保留在显示器上，直到例如用户执行另一个动作（例如，发出播放命令以恢复回放）。可以类似地针对任何其他命令显示符号、徽标等（例如，用于倒带、快进、停止、播放等的符号）。因此可以使用命令接收确认490来传送命令特有的信息。

[0099] 在一些实施例中，可以在接收用户查询或命令之后隐藏语音输入界面484。例如，可以动画显示语音输入界面484使其向下滑动，直到其到达显示器112底部的视图之外。可以在不需要向用户显示进一步信息的情况下隐藏语音输入界面484。例如，对于普通或直接命令（例如，将频道改变到十频道，改变到体育频道，播放，暂停，快进，倒带等），可以在接收确认命令之后立即隐藏语音输入界面484，并可以立即执行关联的任务。尽管本文的各种实施例例示和描述了显示器底边缘或顶边缘处的界面，但应当理解，各个界面中的任一个都可以定位于显示器周围的其他位置。例如，语音输入界面484可以从显示器112的侧边缘，在显示器112的中间或在显示器112的角落等处出现。类似地，可以在显示器上各种不同位置处以各种不同取向布置本文描述的各种其他界面示例。此外，尽管本文描述的各个界面被例示为不透明的，但各个界面的任一个可以是透明的或通过其他方式允许图像（模糊地或全部）被通过界面看到（例如，覆盖媒体内容上的界面内容而不完全使下方媒体内容模糊）。

[0100] 在其他示例中，可以在语音输入界面484之内或在不同界面中显示查询结果。图5示出了视频480上方的示例性媒体内容界面510，具有图4E的转录查询的示例性结果。在一些示例中，虚拟助理查询结果可包括媒体内容以取代文本内容或作为其补充。例如，虚拟助理查询结果可包括电视机节目、视频、音乐等。一些结果可包括立即可以回放的媒体，而其他结果可包括可用于购买等的媒体。

[0101] 如图所示，媒体内容界面510可以比语音输入界面484尺寸更大。在一个实施例中，语音输入界面484可以是更小的第一尺寸以容纳语音输入信息，而媒体内容界面510可以是更大的第二尺寸，以容纳查询结果，查询结果可包括文本、静态图像和移动图像。通过这种方式，用于传送虚拟助理信息的界面可以根据要传送的内容缩放尺寸，由此限制对屏幕占用面积的入侵（例如，最小地遮挡其他内容，例如视频480）。

[0102] 如图所示，媒体内容界面510可包括（作为虚拟助理查询结果）可选择的视频链接512、可选择的文本链接514和附加内容链接513。在一些实施例中，可以通过将焦点、光标等导航到特定元件并利用遥控器（例如，遥控器106）选择它来选择链接。在其他实施例中，可以利用发给虚拟助理的语音命令（例如，观看那场足球赛，显示关于篮球赛的细节等）选择链接。可选择的视频链接512可包括静止或移动图像，可以是可选择的以使关联视频回放。在一个实施例中，可选择的视频链接512可包括播放关联视频内容的视频。在另一个实施例中，可选择的视频链接512可包括电视机频道的实况馈送。例如，作为关于电视机上当前体育赛事的虚拟助理查询结果，可选择的视频链接512可包括体育频道上足球赛的实况馈送。可选择的视频链接512还可包括任何其他视频、动画、图像等（例如，三角形播放符号）。此外，链接512可以链接到诸如电影、电视机节目、体育赛事、音乐等的任何类型的媒体内容。

[0103] 可选择的文本链接514可包括与可选择的视频链接512相关联的文本内容或者可包括虚拟助理查询结果的文本表示。在一个示例中，可选择的文本链接514可包括对从虚拟助理查询得到的媒体的描述。例如，可选择的文本链接514可包括电视机节目的名称、电影标题、对体育赛事的描述、电视机频道名称或号码等。在一个实施例中，选择文本链接514可以使关联媒体内容回放。在另一个实施例中，选择文本链接514可以提供关于媒体内容或其

他虚拟助理查询结果的附加详细信息。附加内容链接513可以链接到虚拟助理查询的附加结果并导致其显示。

[0104] 尽管图5中示出了特定媒体内容的示例,但应当理解,可包括任何类型的媒体内容作为针对媒体内容的虚拟助理查询结果。例如,可以作为虚拟助理的结果而返回的媒体内容可包括视频、电视机节目、音乐、电视机频道等。此外,在一些实施例中,可以在本文的任何界面中提供类别过滤器以允许用户过滤搜索或查询结果或显示的媒体选项。例如,可以提供可选择的过滤器以根据类型(例如,电影、音乐专辑、书籍、电视机节目等)过滤结果。在其他实施例中,可选择的过滤器可包括流派或内容描述符(例如,喜剧、访谈、特定节目等)。在其他实施例中,可选择的过滤器可包括时间(例如,这周、上周、去年等)。应当理解,可以在本文描述的各种界面的任一种中提供过滤器,以允许用户基于与所显示的内容相关的类别过滤结果(例如,在媒体结果具有各种类型时根据类型过滤,在媒体结果具有各种流派时根据流派过滤,在媒体结果具有各种时间时根据时间过滤,等等)。

[0105] 在其他实施例中,媒体内容界面510可包括除媒体内容结果之外的查询改述。例如,可以在媒体内容结果上方(在可选择的视频链接512和可选择的文本链接514上方)显示用户查询的改述。在图5的示例中,此类用户查询改述可包括如下内容:“这里是现在正在进行的一些体育赛事。”可以类似地显示介绍媒体内容结果的其他文本。

[0106] 在一些实施例中,在显示任何界面之后,包括界面510,用户可以利用新查询(可与先前查询相关或不相关)发起捕获附加语音输入。用户查询可包括对界面元件作用的命令,例如选择视频链接512的命令。在另一个实施例中,用户语音可包括与所显示的内容相关联的查询,例如所显示的菜单信息、播放视频(例如,视频480)等。可基于显示的信息(例如,显示的文本)和/或与所显示的内容相关联的元数据(例如,与播放视频相关联的元数据)针对此类查询确定响应。例如,用户可以询问界面(例如,界面510)中显示的媒体结果,并可以搜索与该媒体相关联的元数据,以提供答案或结果。然后可以在另一个界面中或同一界面之内提供此类答案或结果(例如,在本文论述的任何界面中)。

[0107] 如上所述,在一个实施例中,可以响应于对文本链接514的选择显示关于媒体内容的附加详细信息。图6A和图6B示出了在选择文本链接514之后,视频480上方的示例性媒体详情界面618。在一个实施例中,在提供附加详细信息时,可以将媒体内容界面510扩展到媒体详情界面618,如图6A的界面扩展转换616所示。具体地讲,如图6A所示,可以在尺寸上扩展所选择的内容,并通过在显示器112上向上扩展界面来提供附加文本信息,以占据更多屏幕占用面积。可以扩展该界面以容纳用户希望的附加详细信息。通过这种方式,界面的尺寸可以随着用户希望的内容量而缩放,由此使得屏幕占用面积入侵最小化,同时仍然传送期望的内容。

[0108] 图6B示出了完全扩展之后的详情界面618。如图所示,详情界面618可以比媒体内容界面510或语音输入界面484具有更大尺寸,以容纳期望的详细信息。详情界面618可包括详细媒体信息622,详细媒体信息622可包括与媒体内容或虚拟助理查询的另一结果相关联的各种详细信息。详细媒体信息622可包括节目标题、节目描述、节目放送时间、频道、剧集概要、电影描述、演员姓名、角色姓名、体育赛事参与者、出品人姓名、导演姓名或与虚拟助理查询结果相关联的任何其他详细信息。

[0109] 在一个实施例中,详情界面618可包括可选择的视频链接620(或者播放媒体内容

的另一个链接),可选择的视频链接620可包括对应的可选择视频链接512的更大版本。这样一来,可选择的视频链接620可包括静止或移动图像,并且可以是可选择的以使关联视频回放。可选择的视频链接620可包括关联视频内容的播放视频、电视机频道的实况馈送(例如,体育频道的足球赛实况馈送)等。可选择的视频链接620还可包括任何其他视频、动画、图像等(例如,三角形播放符号)。

[0110] 如上所述,可以响应于选择视频链接,例如视频链接620或视频链接512而播放视频。图7A和图7B示出了可以响应于选择视频链接(或播放视频内容的其他命令)而显示的示例性媒体转变界面。如图所示,可以利用视频726替代视频480。在一个实施例中,可以扩展视频726以赶超或覆盖视频480,如图7A中的界面扩展转变724所示。该转变的结果可包括图7B的扩展的媒体界面728。如其他界面那样,扩展媒体界面728的尺寸可以足够大,以为用户提供期望的信息;这里,该信息可包括扩展以填充显示器112。扩展媒体界面728于是可以比任何其他界面更大,因为期望的信息可包括整个显示器上的播放媒体内容。尽管未示出,但在一些实施例中,可以在视频726上简要叠加描述性信息(例如,沿着屏幕底部)。此类描述性信息可包括关联节目、视频、频道等的名称。然后可以从视图(例如,在几秒之后)隐藏描述性信息。

[0111] 图8A到图8B示出了示例性语音输入界面836,其可以显示于显示器112上以向用户传送语音输入信息。在一个实施例中,可以在菜单830上方示出语音输入界面836。菜单830可包括各种媒体选项832,语音输入界面836可以类似地显示于任何其他类型菜单(例如,内容菜单、类别菜单、控制菜单、设置菜单、程序菜单等)上方。在一个实施例中,语音输入界面836可被配置为占据显示器112的屏幕占用面积的较大量。例如,语音输入界面836可以比上述语音输入界面484更大。在一个实施例中,可基于背景内容来确定要使用的语音输入界面的尺寸(例如,更小的界面484或更大的界面836)。在背景内容包括移动图像时,例如,可以显示较小尺寸的语音输入界面(例如,界面484)。另一方面,在背景内容包括静态图像(例如,暂停的视频)或菜单时,例如,可以显示大尺寸的语音输入界面(例如,界面836)。通过这种方式,如果用户正在观看视频内容,可以显示较小的语音输入界面,其仅最小地侵占屏幕占用面积;而如果用户正在导航菜单或观看暂停的视频或其他静态图像,可以显示更大的语音输入界面,其能够通过占据附加占用面积而传送更多信息或具有更深入的效果。本文论述的其他界面能够类似地基于背景内容设定不同尺寸。

[0112] 如上所述,可以触发虚拟助理以侦听包含命令或查询的语音输入(或开始记录用于后续处理的语音输入,或开始实时处理语音输入)。可以通过各种方式触发侦听,包括指示,例如用户按压遥控器106上的物理按钮、用户按压用户设备102上的物理按钮、用户按压用户设备102上的虚拟按钮、用户发出可以由始终侦听的设备识别的短语(例如,发出“你好,助理”以开始侦听命令)、用户执行可以由传感器检测到的手势(例如,在相机前方监测)等。在其他实施例中,用户可以按压并保持遥控器106或用户设备102上的物理按钮以发起侦听。在其他实施例中,用户可以按压和保持遥控器106或用户设备102上的物理按钮,同时讲出查询或命令,并能够在完成时释放按钮。可以类似地接收各种其他指示以发起从用户接收语音输入。

[0113] 响应于接收到侦听语音输入的指示,可以在菜单830上方显示语音输入界面836。图8A示出了从显示器112的底部部分向上扩大的大通知区域834。可以在接收到侦听语音输入

入的指示时在大通知区域834中显示语音输入界面836,可以动画显示该界面,以从图示的显示器112的观看区域的底边缘向上滑动。在一些实施例中,在显示交叠的界面时(例如,响应于接收到侦听语音输入的指示),可以缩小和/或在z方向上(如同向显示器112中进一步移动)向后移动背景菜单、暂停视频、静态图像或其他背景内容。背景界面收缩转变831和关联的内指箭头示出了可以如何收缩背景内容(例如,菜单830)——缩小显示的菜单、图像、文本等。这样可以提供背景内容似乎从用户离开,到达新前景界面(例如,界面836)之外的视觉效果。图8B示出了包括菜单830的收缩(缩小)版本的收缩背景界面833。如图所示,收缩的背景界面833(可包括边界)可以在将焦点转放到前景界面836的同时,看起来从用户进一步离开。本文所述其他实施例的任一个中的背景内容(包括背景视频内容)可以在显示交叠界面时类似地收缩和/或在z方向上向后移动。

[0114] 图8B示出了向上滑动到视图之后的语音输入界面836。如上所述,可以在接收语音输入的同时显示各种确认。尽管这里未示出,但语音输入界面836可以通过与上文分别参考图4B、图4C和图4D所述的语音输入界面484以类似方式,类似地显示准备就绪确认486、侦听确认487和/或处理确认488的更大版本。

[0115] 如图8B所示,可以显示命令接收确认838(如上述更小尺寸的命令接收确认490那样)以确认该系统接收并解释了语音输入。转录840也可以被显示并可包括对所接收语音输入的转录(例如,“纽约天气怎么样?”)。在一些实施例中,可以动画显示抄本840以从显示器112的底部向上滑动,可以在图8B中所示的位置短暂显示抄本(例如,几秒钟),然后可以在从视图中消失之前将其滑动到语音输入界面836顶部(例如,如同文本向上卷动并最终到视图之外那样)。在其他实施例中,可以不显示抄本,可以处理用户的命令或查询,并可以执行关联的任务而不显示抄本。

[0116] 在其他实施例中,可以在用户讲话时实时地执行语音转录。在转录文字时,可以在语音输入界面836中显示它们。例如,可与上述侦听确认487的更大版本并排显示文字。在用户完成讲话之后,可以在执行与用户的命令相关联的任务之前简短地显示命令接收确认838。

[0117] 此外,在其他实施例中,命令接收确认838可以传送关于已接收和理解的命令的信息。例如,对于调谐到特定频道的简单请求,可以在调谐频道时将与频道相关联的徽标或数字简短显示为命令接收确认838(例如,几秒钟)。在另一个实施例中,对于选择所显示菜单项(例如,菜单选项832之一)的请求,可以将与所选择的菜单项相关联的图像显示为命令接收确认838。因此可以使用命令接收确认838来传送命令特有的信息。

[0118] 在一些实施例中,可以在接收用户查询或命令之后隐藏语音输入界面836。例如,可以动画显示语音输入界面836使其向下滑动,直到其到达显示器112底部的视图之外。可以在不需要向用户显示更多信息的情况下隐藏语音输入界面836。例如,对于普通或直接命令(例如,将频道改变到十频道,改变到体育频道,播放该电影等),可以在接收确认命令之后立即隐藏语音输入界面836,并可以立即执行关联的任务。

[0119] 在其他实施例中,可以在语音输入界面836之内或在不同界面中显示查询结果。图9示出了菜单830上方的示例性虚拟助理结果界面942(尤其是在收缩的背景界面833上方),具有图8B的转录查询的示例性结果。在一些实施例中,虚拟助理查询结果可包括文本答案,例如文本答案944。虚拟助理查询结果还可包括解决用户查询的媒体内容,例如与可选择的

视频链接946和购买链接948相关联的内容。具体地讲，在该实施例中，用户可以询问纽约指定位置的天气信息。虚拟助理可以提供直接回答用户查询的文本答案944（例如，指示天气看起来很好，并提供温度信息）。作为文本答案944的替代或补充，虚拟助理可与购买链接948和关联文本一起，提供可选择的视频链接946。与链接946和948相关联的媒体还可以提供对用户查询的响应。这里，与链接946和948相关联的媒体可包括指定位置处天气信息的十分钟剪辑——具体地讲，来自称为天气预报频道的电视机频道的纽约的五天预报。

[0120] 在一个实施例中，解决用户查询的该剪辑可包括先前放送的内容的带时间提示的部分（可以从记录或从流服务获得）。在一个实施例中，虚拟助理可基于与语音输入相关联的用户意图，并通过搜索关于可用媒体内容的详细信息，标识此类内容（例如，包括针对所记录节目的元数据，连同详细定时信息或关于流内容的详细信息）。在一些实施例中，用户可能无法访问或订阅特定内容。在此类情况下，可以提供内容供购买，例如经由购买链接948购买。内容的成本可以在选择购买链接948或视频链接946时自动从用户账户扣除或向用户账户收取。

[0121] 图10示出了用于使用虚拟助理来控制电视机交互和使用不同界面来显示关联信息的示例性过程1000。在方框1002，可以从用户接收语音输入。例如，可以在系统100的用户设备102或遥控器106处接收语音输入。在一些实施例中，可以向服务器系统110和/或电视机顶盒104传输并由其接收语音输入（或语音输入的一些或全部的数据表示）。响应于用户发起接收语音输入，可以在显示器（例如，显示器112）上显示各种通知。例如，可以如上文参考图4A-图4E所述，显示准备就绪确认、侦听确认、处理确认和/或命令接收确认。此外，可以对接收的用户语音输入进行转录，并可以显示转录。

[0122] 再次参考图10的过程1000，在方框1004处，可基于语音输入来确定媒体内容。例如，可以确定指向虚拟助理的解决用户查询的媒体内容（例如，通过搜索可用媒体内容等）。例如，可以确定与图4E的转录492相关的媒体内容（“现在正在进行什么体育赛事？”）。此类媒体内容可包括正在用户可以观看的一个或多个电视机频道上演出的实况体育赛事。

[0123] 在方框1006，可以显示具有可选择媒体链接的第一尺寸的第一用户界面。例如，可以如图5所示，在显示器112上显示具有可选择视频链接512和可选择文本链接514的媒体内容界面510。如上所述，媒体内容界面510可以具有更小尺寸以避免干扰背景视频内容。

[0124] 在方框1008，可以接收链接之一的选择。例如，可以接收链接512和/或链接514之一的选择。在方框1010，可以显示具有和选择相关联的媒体内容的更大第二尺寸的第二用户界面。例如，可以如图6B所示，显示具有可选择视频链接620和详细媒体信息622的详情界面618。如上所述，详情界面618可以是更大尺寸，以传送期望的附加详细媒体信息。类似地，在选择视频链接620时，可以如图7B所示显示带有视频726的扩展媒体界面728。如上所述，扩展媒体界面728可以仍然是更大尺寸，以向用户提供期望的媒体内容。通过这种方式，可以设定本文论述的各个界面的尺寸，以容纳期望的内容（包括扩展成更大尺寸的界面或收缩到更小尺寸的界面），同时以其他方式占据有限的屏幕占用面积。因此可以使用过程1000使用虚拟助理来控制电视机交互并利用不同界面来显示关联信息。

[0125] 在另一个实施例中，可以在控制菜单上方而不是背景视频内容上方显示更大尺寸的界面。例如，可以如图8B所示在菜单830上方显示语音输入界面836，可以如图9所示在菜单830上方显示助理结果界面942，而可以如图5所示在视频480上方显示更小的媒体内容界

面510。通过这种方式,可以至少部分通过背景内容的类型确定界面的尺寸(例如,界面占据的屏幕占用面积的量)。

[0126] 图11示出了用户设备102上的示例性电视机媒体内容,用户设备可包括具有触摸屏246(或另一种显示器)的移动电话、平板电脑、遥控器等。图11示出了包括具有多个电视机节目1152的电视机列表的界面1150。界面1150例如可以对应于用户设备102上的特定应用,例如,电视机控制应用、电视机内容列表应用、互联网应用等。在一些实施例中,可以使用用户设备102上(例如,触摸屏246上)显示的内容从与该内容相关的语音输入确定用户意图,可以使用用户意图使得在另一个设备和显示器(例如,在电视机顶盒104和显示器112上和/或扬声器111上)上回放或显示内容。例如,可以使用用户设备102上的界面1150中显示的内容消除用户请求的歧义并从语音输入确定用户意图,然后可以使用确定的用户意图经由电视机顶盒104播放或显示媒体。

[0127] 图12示出了使用虚拟助理的示例性电视机控制。图12示出了界面1254,其可包括格式化为助理和用户之间的会话对话的虚拟助理界面。例如,界面1254可包括提示用户提出请求的助理问候1256。然后可以转录随后接收的用户语音,例如转录的用户语音1258,来回显示会话。在一些实施例中,界面1254可以响应于触发发起接收语音输入(诸如按钮按压、关键短语等触发)而出现于用户设备102上。

[0128] 在一个实施例中,经由电视机顶盒104(例如,在显示器112和扬声器111上)播放内容的用户请求可能包括对用户设备102上显示的某项内容的歧义援引。例如,转录的用户语音1258包括对“那场”足球赛的援引(“显示那场足球赛。”)。单独从语音输入来讲,期望的特定足球赛可能并不清楚。然而,在一些实施例中,可使用在用户设备102上显示的内容消除用户请求的歧义并确定用户意图。在一个实施例中,用户提出请求之前(例如,界面1254出现于触摸屏246上之前)在用户设备102上显示的内容可用于确定用户意图(内容可能出现于界面1254之内,例如先前的查询和结果)。在例示的实施例中,可以使用图11的界面1150中显示的内容从命令确定用户意图以展示“那场”足球赛。电视机节目1152的电视机列表包括各种不同的节目,其中之一的标题是频道5上出现的“足球”。可以使用足球列表的出现从讲出“那场”足球赛确定用户的意图。具体地讲,用户提到“那场”足球赛可以被解析成出现于界面1150的电视机列表中的足球节目。因此,虚拟助理可以使得回放用户期望的特定足球赛(例如,通过使得电视机顶盒104调谐到适当频道并显示该比赛)。

[0129] 在其他实施例中,用户可以通过各种其他方式援引界面1150中所示的电视机节目(例如,八频道上的节目、新闻、戏剧节目、广告、首映式等),可以类似地基于所显示的内容来确定用户意图。应当理解,还可以进一步结合所显示的内容使用与所显示的内容相关联的元数据(例如,电视机节目描述)、模糊匹配技术、同义匹配等,确定用户意图。例如,可以将术语“广告”匹配到描述“付费节目”(例如,使用同义和/或模糊匹配技术)以从显示“广告”的请求确定用户意图。类似地,可在确定用户意图时分析对特定电视机节目的描述。例如,可以在法庭戏剧的详细描述中识别术语“法律”,并可以从观看“法律”节目的用户请求,基于与界面1150中显示的内容相关联的详细描述,确定用户意图。因此可以使用显示的内容和与其相关联的数据消除用户请求的歧义并确定用户意图。

[0130] 图13示出了用户设备102上的示例性图画和视频内容,用户设备可包括具有触摸屏246(或另一种显示器)的移动电话、平板电脑、遥控器等。图13示出了包括照片和视频列

表的界面1360。界面1360例如可以对应于用户设备102上的特定应用，例如，媒体内容应用、文件导航应用、存储应用、远程存储管理应用、相机应用等。如图所示，界面1360可包括视频1362、相册1364（例如，多个图片的组）和照片1366。如上文参考图11和图12所述，可使用在用户设备102上显示的内容从与该内容相关的语音输入确定用户意图。然后可以使用用户意图使得在另一设备和显示器上回放或显示内容（例如，在电视机顶盒104和显示器112上和/或扬声器111上）。例如，可以使用用户设备102上的界面1360中显示的内容消除用户请求的歧义并从语音输入确定用户意图，然后可以使用确定的用户意图经由电视机顶盒104播放或显示媒体。

[0131] 图14示出了使用虚拟助理的示例性媒体显示控制。图14示出了界面1254，其可包括格式化为助理和用户之间的会话对话的虚拟助理界面。如图所示，界面1254可包括提示用户提出请求的助理问候1256。在对话之内，然后可以如图14的示例所示转录用户语音。在一些实施例中，界面1254可以响应于触发发起接收语音输入（诸如按钮按压、关键短语等触发）而出现于用户设备102上。

[0132] 在一个实施例中，经由电视机顶盒104（例如，在显示器112和扬声器111上）播放媒体内容或显示媒体的用户请求可能包括对用户设备102上显示的某项内容的歧义援引。例如，转录的用户语音1468包括对“那段”视频的援引（“显示那段视频。”）。单独从语音输入来讲，援引的特定视频可能并不清楚。然而，在一些实施例中，可使用在用户设备102上显示的内容消除用户请求的歧义并确定用户意图。在一个实施例中，用户提出请求之前（例如，界面1254出现于触摸屏246上之前）在用户设备120上显示的内容可用于确定用户意图（内容可能出现于界面1254之内，例如先前的查询和结果）。在用户语音1468的实施例中，可以使用图13的界面1360中显示的内容从命令确定用户意图以显示“那段”视频。界面1360中的照片和视频列表包括各种不同的照片和视频，包括视频1362、相册1354和照片1366。因为仅有一个视频出现于界面1360中（例如，视频1362），所以可以使用视频1362在界面1360中的出现从讲出“那段”视频确定用户的意图。具体地讲，可以将用户援引“那段”视频解析成界面1360中出现的视频1362（标题为“毕业视频”）。因此，虚拟助理可以使得回放视频1362（例如，通过使得从用户设备102或远程存储装置向电视机顶盒104传输视频1362并使得开始回放）。

[0133] 在另一个实施例中，转录的用户语音1470包括援引“那个”相册（“播放那个相册的幻灯片。”）。单独从语音输入来讲，援引的特定相册可能并不清楚。在用户设备102上显示的内容可以再次用于消除用户请求的歧义。具体地讲，可以使用图13的界面1360中显示的内容从命令确定用户意图以播放“那个”相册的幻灯片。界面1360中的照片和视频的列表包括相册1354。可以使用界面1360中相册1364的出现从讲出“那个”相册来确定用户的意图。具体地讲，可以将用户援引“那个”相册解析成界面1360中出现的相册1364（标题为“毕业相册”）。因此，响应于用户语音1470，虚拟助理可以使得显示幻灯片，包括来自相册1364的照片（例如，通过使得相册1364的照片从用户设备102或远程存储装置传输到电视机顶盒104并使得开始照片幻灯片）。

[0134] 在另一个实施例中，转录的用户语音1472包括援引“最后一张”照片（“在厨房电视机上显示最后一张照片。”）。单独从语音输入来讲，援引的特定照片可能并不清楚。在用户设备102上显示的内容可以再次用于消除用户请求的歧义。具体地讲，可以使用图13的界面

1360中显示的内容从命令确定用户意图以显示“最后一张”照片。界面1360中的照片和视频的列表包括两张单独的照片1366。照片1366在界面1360中出现，尤其是照片1366在界面之内出现的次序，可以用于从讲出“最后一张”照片确定用户的意图。具体地讲，可以将用户援引“最后一张”照片解析为出现于界面1360底部的照片1366（日期为2014年6月21日）。因此，响应于用户语音1472，虚拟助理可以使得显示界面1360中所示的最后一张照片1366（例如，通过导使得最后一张照片1366从用户设备102或远程存储装置传输到电视机顶盒104并使得显示照片）。

[0135] 在其他实施例中，用户可以通过各种其他方式援引界面1360中所示的媒体内容（例如，最后一对照片、所有视频、所有照片、毕业相册、毕业视频、6月21日的照片等），可以类似地基于所显示的内容来确定用户意图。应当理解，还可以进一步结合所显示的内容使用与所显示的内容相关联的元数据（例如，时间戳、位置信息、标题、描述等）、模糊匹配技术、同义匹配等，确定用户意图。因此可以使用所显示的内容和与其相关联的数据消除用户请求的歧义并确定用户意图。

[0136] 应当理解，可以在确定用户意图时使用任何应用的任何应用界面中的任何类型的所显示的内容。例如，可以在语音输入中援引互联网浏览器应用中的网页上显示的图像，并可以分析所显示的网页内容以识别期望的图像。类似地，可以在语音输入中根据标题、流派、艺术家、乐队名等援引音乐应用中音乐列表中的音乐轨道，并可以使用音乐应用中所显示的内容（和某些实施例中的关联元数据）从语音输入确定用户意图。如上所述，然后可以使用所确定的用户意图使得经由另一设备，例如经由电视机顶盒104显示或回放媒体。

[0137] 在一些实施例中，可以采用用户标识、用户认证和/或设备认证确定是否可以许可媒体控制，确定媒体内容是否可以显示，确定访问许可等。例如，可以确定是否授权特定用户设备（例如，用户设备102）控制例如电视机顶盒104上的媒体。可基于注册、配对、信任确定、密码、安全问题、系统设置等授权用户设备。响应于确定特定用户设备已经被授权，可以许可尝试控制电视机顶盒104（例如，可以响应于确定授权请求设备控制媒体来播放媒体内容）。相反，可以忽略来自未授权设备的媒体控制命令或请求和/或可以提示此类设备的用户注册其设备以在控制特定电视机顶盒104时使用。

[0138] 在另一个实施例中，可以识别特定用户，并可使用与该用户相关联的个人数据来确定请求的用户意图。例如，可基于语音输入，例如通过利用用户的声波纹的语音识别来识别用户。在一些实施例中，用户可以讲出特定短语，该短语被分析以进行语音识别。在其他实施例中，可以利用语音识别来分析指向虚拟助理的语音输入请求，以识别讲话者。也可基于（例如，在用户的个人设备102上）语音输入样本的源识别用户。也可基于口令、密码、菜单选择等识别用户。然后可基于所识别用户的个人数据解释从用户接收的语音输入。例如，可基于来自用户的先前请求、用户拥有的媒体内容、用户设备上存储的媒体内容、用户偏好、用户设置、用户人口统计信息（例如，所讲的语言等）、用户概况信息、用户支付方法、或与特定识别的用户相关联的各种其他个人信息，确定语音输入的用户意图。例如，可基于个人数据消除援引收藏列表等的语音输入的歧义，并可以识别用户的个人收藏列表。可以类似地基于用户标识消除援引“我的”照片、“我的”视频、“我的”表演等的语音输入的歧义，以正确识别与所识别用户相关联的照片、视频和表演（例如，存储于个人用户设备上的照片等）。类似地，可以消除请求购买内容的语音输入以确定所识别用户的支付方法应当为购物而付费。

(与另一个用户的支付方法相反)。

[0139] 在一些实施例中,可以使用用户的认证确定是否允许用户访问媒体内容、购买媒体内容等。例如,可以使用语音识别验证特定用户的身份(例如,使用其声波纹),以许可用户利用用户的支付方法进行购买。类似地,可以使用口令等认证用户以许可购物。在另一个实施例中,可以使用语音识别验证特定用户的身份以确定是否许可用户观看特定节目(例如,具有特定父母指导分级的节目、具有特定年龄适应分级的电影等)。例如,可基于指示请求者不是能够观看此类内容的授权用户(例如,父母)的语音识别,拒绝儿童对特定节目的请求。在其他实施例中,可以使用语音识别来确定用户是否能够访问特定的订阅内容(例如,基于语音识别限制对优质频道内容的访问)。在一些实施例中,用户可以讲出特定短语,该短语被分析以进行语音识别。在其他实施例中,可以利用语音识别来分析指向虚拟助理的语音输入请求,以识别讲话者。因此可以响应于第一次确定通过各种方式的任一种授权用户而播放特定的媒体内容。

[0140] 图15示出了在移动用户设备和媒体显示设备上具有结果的示例性虚拟助理交互。在一些实施例中,虚拟助理可以提供关于超过一个设备的信息和控制,例如用户设备102以及电视机顶盒104。此外,在一些实施例中,可以使用用于用户设备102上的控制和信息的同一虚拟助理界面发出对控制电视机顶盒104上媒体的请求。这样一来,虚拟助理系统可以确定是否在用户设备102上或在电视机顶盒104上显示结果或执行任务。在一些实施例中,在采用用户设备102控制电视机顶盒104时,可以通过在用户设备102上(例如,在触摸屏246上)显示信息来使虚拟助理界面在与电视机顶盒104相关联的显示器(例如,显示器112)上的侵入最小化。在其他实施例中,可以单独在显示器112上显示虚拟助理信息,或者可以在用户设备102和显示器112两者上显示虚拟助理信息。

[0141] 在一些实施例中,可以确定是否应当直接在用户设备102上或在与电视机顶盒104相关联的显示器112上显示虚拟助理查询结果。在一个实施例中,响应于确定查询的用户意图包括对信息的请求,可以在用户设备102上显示信息响应。在另一个实施例中,响应于确定查询的用户意图包括用于播放媒体内容的请求,可以经由电视机顶盒104播放响应于查询的媒体内容。

[0142] 图15示出了在虚拟助理和用户之间具有会话对话示例的虚拟助理界面1254。助理问候1256可以提示用户提出请求。在第一查询中,转录的用户语音1574(也可以通过其他方式键入或输入)包括对与所显示媒体内容相关联的信息答案的请求。具体地讲,转录的用户语音1574查询谁正在参加可能例如在用户设备102上的界面上(例如,图11的界面1150中列出)或在显示器112上(例如,在图5的界面510中列出或作为视频726在图7B的显示器112上播放)显示的足球赛。可基于所显示的媒体内容来确定转录用户语音1574的用户意图。例如,可基于在用户设备102或显示器112上显示的内容识别正在论述的特定足球赛。转录得用户语音1574的用户意图可包括获得基于所显示的内容详述参与所识别的足球赛中的队伍的信息答案。响应于确定用户意图包括对信息答案的请求,该系统可以确定在图15的界面1254之内显示响应(与显示器112相反)。在一些实施例中,可基于与所显示的内容相关联的元数据(例如,基于对电视机列表中的足球赛的描述)确定对查询的响应。如图所示,助理响应1576因此可以被显示于界面1254中的用户设备102的触摸屏246上,标识Alpha和Zeta队正在参加比赛。因此,在一些实施例中,可基于确定该查询包括信息请求在用户设备102

上的界面1254之内显示信息响应。

[0143] 然而,界面1254中的第二查询包括媒体请求。具体地讲,转录的用户语音1578请求将显示的媒体内容改变为“比赛”。可基于显示的内容来确定转录的用户语音1578的用户意图(例如,识别用户希望看哪场比赛),例如,图5的界面510中列出的比赛,图11的界面1150中列出的比赛,先前查询中援引的比赛(例如,在转录的用户语音1574中)等。转录的用户语音1578的用户意图因此可包括将显示的内容改变为特定比赛——这里,是Alpha和Zeta队的足球赛。在一个实施例中,可以在用户设备102上显示比赛。然而,在其他实施例中,基于包括播放媒体内容的请求的查询,可以经由电视机顶盒104显示比赛。具体地讲,响应于确定用户意图包括用于播放媒体内容的请求,该系统可以确定在显示器112上经由电视机顶盒104显示媒体内容结果(与图15中的界面1254之内相反)。在一些实施例中,可以在界面1254中或显示器112上显示确认虚拟助理的期望动作的响应或改述(例如,“更改到足球赛。”)。

[0144] 图16示出了在媒体显示设备和移动用户设备上具有媒体结果的示例性虚拟助理交互。在一些实施例中,虚拟助理可以提供对用户设备102和电视机顶盒104上媒体的访问。此外,在一些实施例中,可以使用用于用户设备102上媒体的同一虚拟助理界面发出对电视机顶盒104上媒体的请求。这样一来,虚拟助理系统可以确定是否在用户设备102上或经由电视机顶盒104在显示器112上显示媒体结果。

[0145] 在一些实施例中,可基于媒体结果格式、用户偏好、默认设置、请求自身中的明确命令等确定是否在设备102上或显示器112上显示媒体。例如,可以使用查询的媒体结果的格式确定默认在哪个设备上显示媒体结果(例如,无需特定指令)。电视机节目可以更适合在电视机上显示,更大格式的视频可以更适合在电视机上显示,缩略图照片可以更适合在用户设备上显示,小格式的Web视频可以更适合在用户设备上显示,各种其他媒体格式可以更适合在较大电视机屏幕上或较小用户设备显示器上显示。因此,响应于(例如,基于媒体格式)确定应当在特定显示器上显示媒体内容,可以默认在该特定显示器上显示媒体内容。

[0146] 图16示出了虚拟助理界面1254,带有与播放或显示媒体内容相关的查询的示例。助理问候1256可以提示用户提出请求。在第一查询中,转录的用户语音1680包括显示足球赛的请求。如上述实施例中那样,可基于显示的内容来确定转录的用户语音1680的用户意图(例如,识别用户希望看哪场比赛),例如,图5的界面510中列出的比赛,图11的界面1150中列出的比赛,先前查询中援引的比赛等。转录的用户语音1680的用户意图因此可包括显示特定足球赛,该足球赛例如在电视机上被放送。响应于确定用户意图包括显示针对电视机而格式化的媒体的请求(例如,电视机放映的足球赛),该系统可以自动确定经由电视机顶盒104在显示器112上显示期望的媒体(与用户设备102自身相反)。虚拟助理系统然后可以使电视机顶盒104调谐到足球赛,并在显示器112上显示它(例如,通过执行必要的任务和/或发送适当的命令)。

[0147] 然而,在第二查询中,转录的用户语音1682包括显示球队队员图片(例如,“Alpha队的图片”)的请求。如在上述实施例中那样,可以确定转录的用户语音1682的用户意图。转录的用户语音1682的用户意图可包括对与“Alpha队”相关联的图片进行搜索(例如,网络搜索)并显示所得的图片。响应于确定用户意图包括显示可以按照缩略图格式呈现的媒体或与网络搜索相关联的媒体或没有特定格式的其他非特定媒体的请求,该系统可以自动决定

在用户设备102的界面1254中的触摸屏246上显示期望的媒体结果(与经由电视机顶盒104在显示器112上显示所得的图片相反)。例如,如图所示,可以响应于用户的查询在用户设备102上的界面1254之内显示缩略图照片1684。因此,虚拟助理系统可以使得默认在用户设备102上显示特定格式的媒体或可以通过特定格式呈现的媒体(例如,在一组缩略图中)。

[0148] 应当理解,在一些实施例中,可以在用户设备102上显示用户语音1680中援引的足球赛,并可以经由电视机顶盒104在显示器112上显示照片1684。然而,可基于媒体格式自动确定默认显示设备,由此简化用于用户的媒体命令。在其他实施例中,可基于用户偏好、默认设置、最近用于显示内容的设备、识别用户和与用户相关联的设备的语音识别等,确定用于显示请求的媒体内容的默认设备。例如,用户可以设置偏好或者可以设置默认配置以经由电视机顶盒104在显示器112上显示特定类型的内容(例如,视频、幻灯片、电视机节目等)并在用户设备102的触摸屏246上显示其他类型的内容(例如,缩略图、照片、网络视频等)。类似地,可以设置偏好或默认配置以通过在一个设备或另一个上显示内容而对特定查询做出响应。在另一个实施例中,可以在用户设备102上显示所有内容,除非用户指示不这样做。

[0149] 在其他实施例中,用户查询可包括在特定显示器上显示内容的命令。例如,图14的用户语音1472包括在厨房电视机上显示照片的命令。因此,该系统可以使得在与用户厨房相关联的电视机显示器上显示照片,与在用户设备102上显示照片相反。在其他实施例中,用户可以通过各种其他方式指示使用哪个显示设备(例如,在电视机上,在大屏幕上,在起居室中,在卧室中,在我的平板电脑上,在我的电话上等)。因此可以通过各种不同方式确定用于显示虚拟助理查询的媒体内容结果的显示设备。

[0150] 图17示出了基于接近性的示例性媒体设备控制。在一些实施例中,用户可能在同一家庭之内或在同一网络上具有多个电视机和电视机顶盒。例如,家庭可以在起居室中具有电视机和机顶盒,在卧室中有另一组,在厨房中有另一组。在其他实施例中,多个机顶盒可以连接到同一网络,例如公寓或办公楼中的公共网络。尽管用户可以对用于特定机顶盒的遥控器106和用户设备102进行配对、连接或通过其他方式授权,以避免未授权的访问,但在其他实施例中,可以使用遥控器和/或用户设备控制超过一个机顶盒。用户可以例如使用单个用户设备102控制卧室、起居室和厨房中的机顶盒。用户还可以例如使用单个用户设备102控制其自己公寓中的自己的机顶盒,以及控制邻居公寓中的邻居机顶盒(例如,与邻居共享来自用户设备102的内容,例如在邻居的电视机上显示用户设备102上存储的照片的幻灯片)。因为用户可以使用单个用户设备102控制多个不同的机顶盒,所以该系统可以确定向多个机顶盒中的哪个机顶盒发送命令。同样,因为家庭可以具有能够操作多个机顶盒的多个遥控器106,所以该系统可以类似地确定向多个机顶盒中的哪个机顶盒发送命令。

[0151] 在一个实施例中,可以使用设备的接近性确定向多个机顶盒中的哪个发送命令(或在哪个显示器上显示所请求的媒体内容)。可以在用户设备102或遥控器106和多个机顶盒的每个之间确定接近性。然后可以将发出的命令发送到最近的机顶盒(或可以在最近的显示器上显示所请求的媒体内容)。可以通过各种方式的任一种,例如飞行时间测量(例如,使用射频)、蓝牙LE、电子查验信号、接近传感器、声音行进测量等,确定接近性。然后可以比较测量或近似的距离,可以向最近距离的设备发出命令(例如,最近的机顶盒)。

[0152] 图17示出了多设备系统1790,包括具有第一显示器1786的第一机顶盒1792和具有第二显示器1788的第二机顶盒1794。在一个实施例中,用户可以从用户设备102发出显示媒

体内容的命令(例如,不必指定在哪里或在哪个设备上)。然后可以确定(或近似确定)与第一机顶盒1792的距离1795和与第二机顶盒1794的距离1796。如图所示,距离1796可以大于距离1795。基于接近性,可以向作为最近设备并最可能匹配用户意图的第一机顶盒1792发出来自用户设备102的命令。在一些实施例中,单个遥控器106也可以用于控制超过一个机顶盒。可基于接近性确定在给定时间用于控制的期望设备。然后可以确定(或近似确定)与第二机顶盒1794的距离1797和与第一机顶盒1792的距离1798。如图所示,距离1798可以大于距离1797。基于接近性,可以向作为最近设备并最可能匹配用户意图的第二机顶盒1794发出来自遥控器106的命令。可以定期地或利用每条例如适应用户进入不同房间并希望控制不同设备的命令,刷新距离测量结果。

[0153] 应当理解,用户可以为命令指定不同设备,在一些情况下,超驰接近性。例如,可以在用户设备102上显示可用显示设备的列表(例如,通过设置名称、指定房间等列出第一显示器1786和第二显示器1788,或者通过设置名称、指定房间等列出第一机顶盒1792和第二机顶盒1794)。用户可以从列表选择设备之一,然后可以向所选择的设备发送命令。然后可以通过在所选择的设备上显示期望的媒体来处理对在用户设备102处发出的媒体内容的请求。在其他实施例中,用户可以讲出期望的设备作为口语命令的部分(例如,在厨房电视机上显示比赛,改变到起居室中的卡通频道等)。

[0154] 在其他实施例中,可基于与特定设备相关联的状态信息确定用于显示所请求媒体内容的默认设备。例如,可以确定耳机(或耳麦)是否附接到用户设备102。响应于确定在接收显示媒体内容的请求时耳机附接到用户设备102,可以默认在用户设备102上显示所请求的内容(例如,假设用户正在消费用户设备102而非电视机上的内容)。响应于确定在接收显示媒体内容的请求时耳机未附接到用户设备102,可以根据本文论述的各种确定方法的任一种在用户设备102或电视机上显示所请求的内容。可以类似地使用其他设备状态信息以确定是否应当在用户设备102或机顶盒104上显示所请求的内容,例如用户设备102或机顶盒104周围的环境照明、其他设备与用户设备102或机顶盒104的接近性、用户设备102的取向(例如,横向取向可以更可能指示用户设备102上的期望观察)、机顶盒104的显示状态(例如,在睡眠模式中)、上次在特定设备上交互之后过去的时间、或针对用户设备102和/或机顶盒104的各种其他状态指示符的任一种。

[0155] 图18示出了用于使用虚拟助理和多个用户设备来控制电视机交互的示例性过程1800。在方框1802,可以在具有第一显示器的第一设备处从用户接收语音输入。例如,可以在系统100的用户设备102或遥控器106处从用户接收语音输入。第一显示器可包括用户设备102的触摸屏246或与一些实施例中的遥控器106相关联的显示器。

[0156] 在方框1804,可基于在第一显示器上所显示的内容从语音输入确定用户意图。例如,可以分析诸如图11的界面1150中的电视机节目或图13的界面1360中的照片和视频的内容,并用于确定针对语音输入的用户意图。在一些实施例中,用户可以通过歧义方式援引在第一显示器上所显示的内容,可以通过分析在第一显示器上所显示的内容以解析援引来消除援引的歧义(例如,确定针对“那段”视频、“那个”相册、“那场”比赛等的用户意图),如上文参考图12和图14所述。

[0157] 再次参考图18的过程1800,在方框1806处,可基于用户意图来确定媒体内容。例如,可基于用户意图识别特定视频、照片、相册、电视机节目、体育赛事、音乐轨道等。在上文

所述的图11和图12的示例中,例如,可基于援引图11的界面1150中所示的“那场”足球赛的用户意图标识五频道上显示的特定足球赛。在上文所述的图13和图14的示例中,可基于从图14的语音输入示例确定的用户意图,识别标题为“毕业视频”的特定视频1362、标题为“毕业相册”的特定相册1364或特定照片1366。

[0158] 再次参考图18的过程1800,在方框1808处,可以在与第二显示器相关联的第二设备上播放媒体内容。例如,可以经由电视机顶盒104在具有扬声器111的显示器112上播放所确定的媒体内容。播放媒体内容可包括在电视机顶盒104或另一个设备上调谐到特定电视机频道,播放特定视频,显示照片的幻灯片,显示特定照片,播放特定音频轨道等。

[0159] 在一些实施例中,可以确定是否应当在与第一设备(例如,用户设备102)相关联的第一显示器上或与第二设备(例如,电视机顶盒104)相关联的第二显示器上显示对指向虚拟助理的语音输入的响应。例如,如上文参考图15和图16所述,可以在用户设备102上显示适合在更小屏幕上显示的信息答案或媒体内容,而可以在与机顶盒104相关联的显示器上显示适合在更大屏幕上显示的媒体响应或媒体内容。如上文参考图17所述,在一些实施例中,用户设备102和多个机顶盒之间的距离可以用于确定在哪个机顶盒上播放媒体内容或向哪个机顶盒发出命令。可以类似地做出各种其他确定以在有多个设备可以交互时提供方便且用户友好的体验。

[0160] 在一些实施例中,由于可以如上所述使用在用户设备102上显示的内容来通知语音输入的解释,因此可同样在使用显示器112上显示的内容来通知语音输入的解释。具体地讲,可将在与电视机顶盒104相关联的显示器上所显示的内容连同与该内容相关联的元数据一起用于从语音输入确定用户意图,消除用户查询的歧义,对内容相关的查询做出响应等。

[0161] 图19示出了示例性语音输入界面484(如上所述),在背景中示出了关于视频480的虚拟助理查询。在一些实施例中,用户查询可包括关于显示器112上显示的媒体内容的问题。例如,转录1916包括请求识别女演员的查询(“那些女演员是谁?”)。在显示器112上显示的内容——连同关于该内容的元数据或其他描述性信息——可用于从与该内容相关的语音输入确定用户意图以及确定对查询的响应(响应包括信息性响应以及向用户提供媒体选择的媒体响应)。例如,视频480、对视频480的描述、视频480的角色和演员列表、视频480的分级信息、视频480的流派信息和与视频480相关联的各种其他描述性信息可用于消除用户请求的歧义并确定对用户查询的响应。关联的元数据可包括,例如,针对角色1910、角色1912和角色1914的标识信息(例如,角色名称连同扮演角色的女演员的姓名)。用于任何其他内容的元数据可以类似地包括标题、描述、角色列表、演员列表、队员列表、流派、出品人姓名、导演姓名或与在显示器上所显示的内容相关联的显示调度或显示器上媒体内容的观看历史(例如,最近显示的媒体)。

[0162] 在一个实施例中,指向虚拟助理的用户查询可包括对显示器112上显示的某物的歧义援引。例如,转录1916包括对“那些”女演员的援引(“那些女演员是谁?”)。单独从语音输入来讲,用户询问的特定女演员可能并不清楚。然而,在一些实施例中,可使用在显示器112上显示的内容和关联的元数据消除用户请求的歧义并确定用户意图。在例示的实施例中,可使用在显示器112上显示的内容从对“那些”女演员的援引确定用户意图。在一个实施例中,电视机顶盒104可以标识播放内容连同与该内容相关联的细节。在这种情况下,电视

机顶盒104可以识别视频480的标题连同各种描述性内容。在其他实施例中,可以示出电视机表演、体育赛事或其他内容,其可以结合关联的元数据使用以确定用户意图。此外,在本文论述的各种实施例的任一个中,语音识别结果和意图确定可以对与所显示的内容相关联的项比替代方式加权更高。例如,可以对屏幕上角色的演员姓名在那些演员出现于屏幕上时(或在正在播放出现他们的演出时)加权更高,这样可以提供与所显示的内容相关联的可能用户请求的精确的语音识别和意图确定。

[0163] 在一个实施例中,可以使用与视频480相关联的角色和/或演员列表识别视频480中出现的所有或大部分主要女演员,这可包括女演员1910、1912和1914。可以将所识别的女演员作为可能结果而返回(如果元数据分辨率粗糙,包括更少或另外的女演员)。然而,在另一个实施例中,与视频480相关联的元数据可包括在给定时间哪些男演员和女演员出现于屏幕上的标识,并且可从元数据来确定在进行查询时出现的女演员(例如,具体标识女演员1910、1912和1914)。在另一个实施例中,可以使用面部识别应用从显示器112上显示的图像识别女演员1910、1912和1914。在其他实施例中,可以使用与视频480相关联的各种其他元数据和各种其他识别方法在推断“那些”女演员时识别用户的可能意图。

[0164] 在一些实施例中,在显示器112上显示的内容可以在提交查询和确定响应期间改变。这样一来,可以使用媒体内容的观看历史来确定用户意图并确定对查询的响应。例如,如果在产生对查询的响应之前视频480移动到另一个视图(例如,具有其他角色),可基于讲出查询时用户的视图确定查询结果(例如,用户发起查询时屏幕上显示的角色)。在一些情况下,用户可能暂停播放媒体以发出查询,可与关联的元数据一起使用暂停时显示的内容来确定用户意图和对查询的响应。

[0165] 给定所确定的用户意图,可以向用户提供查询结果。图20示出了示例性助理响应界面2018,包括助理响应2020,其可包括从图19的转录1916的查询确定的响应。如图所示,助理响应2020可包括,每位女演员的姓名及其在视频480中的关联角色的列表(“女演员Jennifer Jones饰演角色Blanche;女演员Elizabeth Arnold饰演角色Julia;女演员Whitney Davidson饰演角色Melissa。”)。响应2020中列出的女演员和角色可以对应于显示器112上出现的角色1910、1912和1914。如上所述,在一些实施例中,在显示器112上显示的内容可以在提交查询和确定响应期间改变。这样一来,响应2020可包括关于显示器112上可能不再出现的内容或角色的信息。

[0166] 如显示器112上显示的其他界面那样,助理响应界面2018可以占据屏幕占用面积的最小量,同时提供充分大空间以传送期望的信息。在一些实施例中,如显示器112上界面中显示的其他文本那样,助理响应2020可以从显示器112底部向上滚动到图20中所示的位置,显示特定量时间(例如,基于响应长度的延迟),并向上滚动到视图之外。在其他实施例中,界面2018可以在延迟之后向下滑动到视图之外。

[0167] 图21和图22示出了基于在显示器112上显示的内容来确定用户意图并对查询做出响应的另一个示例。图21示出了具有用于与视频480相关联的媒体内容的虚拟助理查询的示例性语音输入界面484。在一些实施例中,用户查询可包括对与显示器112上显示的媒体相关联的媒体内容的请求。例如,用户可基于例如角色、演员、流派等,请求与特定媒体相关联的其他电影、电视机节目、体育赛事等。例如,转录2122包括请求与视频480中的女演员相关联的其他媒体的查询,由视频480中其角色姓名援引(“Blanche在其他什么剧中出现

过?”)。在显示器112上显示的内容——连同关于内容的元数据或其他描述性信息——可以再次用于从与该内容相关的语音输入确定用户意图以及确定对查询的响应(信息性或导致媒体选择)。

[0168] 在一些实施例中,指向虚拟助理的用户查询可包括使用角色姓名、演员姓名、节目名称、队员姓名等的歧义援引。没有在显示器112上显示的内容语境及其关联的元数据,此类援引可能难以精确解析。转录2122例如包括对来自视频480的名为“Blanche”的角色的援引。单独从语音输入来讲,用户询问的特定女演员或其他个人可能并不清楚。然而,在一些实施例中,可使用在显示器112上显示的内容和关联的元数据来消除用户请求的歧义并确定用户意图。在例示的实施例中,可使用在显示器112上显示的内容和关联的元数据从角色名称“Blanche”确定用户意图。在这种情况下,可以使用与视频480相关联的角色列表确定“Blanche”可能是指视频480中的角色“Blanche”。在另一个实施例中,可以使用详细的元数据和/或面部识别确定名称为“Blanche”的角色出现于屏幕上(或在发起用户查询时出现于屏幕上),使得与该角色相关联的女演员成为用户查询的最可能意图。例如,可以确定角色1910、1912和1914出现于显示器112上(或在发起用户查询时出现于显示器112上),然后可以援引他们的关联角色以确定援引角色Blanche的查询的用户意图。然后可以使用演员列表识别饰演Blanche的女演员,可以进行搜索以识别出现所识别的女演员的其他媒体。

[0169] 给定所确定的用户意图(例如,角色援引“Blanche”的分辨率)和对查询结果的确定(例如,与饰演“Blanche”的女演员相关联的其他媒体),可以向用户提供响应。图22示出了示例性助理响应界面2224,包括助理文本响应2226和可选择的视频链接2228,它们可以是响应于图21的转录2122的查询而给出的。如图所示,助理文本响应2226可包括介绍可选择视频链接2228的用户请求的改述。助理文本响应2226还可包括消除用户查询歧义的指示——具体地讲,将女演员Jennifer Jones标识为饰演视频480中的角色Blanche。此类改述可以向用户确认虚拟助理正确地解释了用户查询并提供期望的结果。

[0170] 助理响应界面2224还可包括可选择的视频链接2228。在一些实施例中,可以提供各种媒体内容作为虚拟助理查询结果,包括电影(例如,界面2224的电影A和电影B)。作为查询结果显示的媒体内容可包括可供用户消费的媒体(免费、购买或作为订阅的一部分)。用户可以选择所显示的媒体以观看或消费所得的内容。例如,用户可以选择可选择的视频链接2228之一(例如,使用遥控器、语音命令等)以观看女演员Jennifer Jones出现的其他电影之一。响应于选择了可选择的视频链接2228之一,可以播放与该选择相关联的视频,替代显示器112上的视频480。因此,可以使用显示的媒体内容和关联的元数据从语音输入确定用户意图,在一些实施例中,可以提供可播放的媒体作为结果。

[0171] 应当理解,用户可以在形成查询时援引与所显示的内容相关联的演员、队员、角色、位置、球队、体育赛事细节、电影主题或与所显示的内容相关联的各种其他信息,虚拟助理系统可以类似地消除此类请求的歧义并基于所显示的内容和关联的元数据来确定用户意图。类似地,应当理解,在一些实施例中,结果可包括与查询相关联的媒体建议,例如,与作为查询主题的人相关联的电影、电视机表演或体育赛事(无论用户是否特定请求此类媒体内容)。

[0172] 此外,在一些实施例中,用户查询可包括对与媒体内容自身相关联的信息的请求,例如关于角色、剧集、电影情节、前一场景等的查询。如上述实施例那样,可以使用显示的内

容和关联的元数据从此类查询确定用户意图并确定响应。例如，用户可能请求对角色的描述(例如，“Blanche在这部电影中做了什么?”)。虚拟助理系统然后可以从与所显示的内容相关联的元数据识别关于角色的所请求信息，例如角色描述或人物(例如，“Blanche是一组律师之一，被认为是Hartford的麻烦制造者。”)。类似地，用户可以请求剧集摘要(例如，“上一集中发生了什么?”)，虚拟助理系统可以搜索并提供对剧集描述。

[0173] 在一些实施例中，在显示器112上显示的内容可包括菜单内容，此类菜单内容可以类似地用于确定语音输入的用户意图和对用户查询的响应。图23A-图23B示出了程序菜单830的示例性页面。图23A示出了第一页媒体选项832，图23B示出了第二页媒体选项832(其可包括扩展到单个页面之外的内容列表中的相继下一页)。

[0174] 在一个实施例中，播放内容的用户请求可包括菜单830中对显示器112上显示的某内容的歧义援引。例如，用户观看菜单830可以请求观看“那场”足球赛、“那场”篮球赛、吸尘器广告、法律节目等。单独从语音输入来讲，期望的特定节目可能并不清楚。然而，在一些实施例中，可使用在显示器112上显示的内容消除用户请求的歧义并确定用户意图。在例示的实施例中，菜单830中的媒体选项(在一些实施例中连同与媒体选项相关联的元数据)可用于从包括歧义援引的命令确定用户意图。例如，“那场”足球赛可以被解析成体育频道上的足球赛。“那场”篮球赛可以被解析成大学体育频道上的篮球赛。吸尘器广告可以被解析成付费节目表演(例如，基于与描述吸尘器的节目相关联的元数据)。法律节目可基于与节目相关联的元数据和/或同义匹配、模糊匹配或其他匹配技术被解析成法庭戏剧。显示器112上的菜单830中出现各种媒体选项832从而可用于消除用户请求的歧义。

[0175] 在一些实施例中，可以利用光标、操纵杆、箭头、按钮、手势等导航所显示的菜单。在此类情况下，可以针对所选择的项目显示焦点。例如，可以通过粗体、下划线、带边界的外框、比其他菜单项更大的尺寸、阴影、反光、光晕和/或任何其他特征来显示所选择的项目，以突出选择了哪个菜单项并具有焦点。例如，图23A中所选择的媒体选项2330可以具有像当前所选择的媒体选项那样的焦点，被示为具有大的下划线类型和边界。

[0176] 在一些示实施例中，播放内容或选择菜单项的请求可包括对具有焦点的菜单项的歧义援引。例如，图23A的用户观看菜单830能够请求播放“那场”演出(例如，“播放那场演出”。)。类似地，用户可以请求与具有焦点的菜单项相关联的各种其他命令，例如，播放、删除、隐藏、提醒我观看、录制等。单独从语音输入来讲，期望的特定菜单项或表演可能并不清楚。然而，可使用在显示器112上显示的内容消除用户请求的歧义并确定用户意图。具体地讲，可以使用所选择的菜单选项2330在菜单830中具有焦点的事实来识别援引“那场”演出、没有主题的命令(例如，播放、删除、隐藏等)或援引具有焦点的媒体内容的任何其他歧义命令的任一种的期望媒体主题。因此可以在从语音输入确定用户意图时使用具有焦点的菜单项。

[0177] 如可以用于消除用户请求歧义的媒体内容的观看历史那样(例如，在用户发起请求时但自从传递过去之后而显示的内容)，可以类似地使用先前显示的菜单或搜索结果内容，在例如移动到稍晚菜单或搜索结果内容之后，消除稍晚用户请求的歧义。例如，图23B示出了具有附加媒体选项832的第二页菜单830。用户可以推进到图23B所示的第二页，但重新参考图23A中所示的第一页中示出的内容(例如，图23A中所示的媒体选项832)。例如，尽管已经移动到菜单830的第二页，但用户仍然能够请求观看“那场”足球赛、“那场”篮球赛或法

律节目——所有这些都是最近显示于前一页菜单830上的媒体选项832。此类援引可能有歧义,但可以使用来自第一页菜单830的最近显示的菜单内容来确定用户意图。具体地讲,可以分析图23A的最近显示的媒体选项832来识别歧义示例请求中援引的特定足球赛、篮球赛或法庭戏剧。在一些示实施例中,可基于最近多久显示了内容来偏置结果(例如,相对于更早观看的结果,对最近观看的结果页加权)。通过这种方式,可以使用显示器112上最近显示什么的观看历史来确定用户意图。应当理解,可以使用任何最近显示的内容,例如先前显示的搜索结果、先前显示的节目、先前显示的菜单等。这样可以允许用户向回援引它们更早看到的某个内容,而无需找到并导航到它们看到它的特定视图。

[0178] 在其他实施例中,可以使用显示器112上的菜单或结果列表中所示的各种显示提示消除用户请求的歧义并确定用户意图。图24示出了被分成各别的示例性媒体菜单,其中之一具有焦点(电影)。图24示出了类别界面2440,其可包括分类媒体选项的圆盘式界面,包括电视机选项2442、电影选项2444和音乐选项2446。如图所示,仅部分显示音乐类别,可以将圆盘式界面偏移以显示右方的附加内容(例如,如箭头所示),如同在圆盘中旋转媒体那样。在例示的实施例中,电影类别具有如下划线标题和边界所示的焦点,尽管焦点可以通过各种其他方式的任何一种表示(例如,使得类别更大以比其他类别显得更接近用户,增加光晕等)。

[0179] 在一些实施例中,播放内容或选择菜单项的请求可包括对一组项目(例如类别)中的菜单项的歧义援引。例如,观看类别界面2440的用户可能请求播放足球节目(“播放足球节目。”)。单独从语音输入来讲,期望的特定菜单项或表演可能并不清楚。此外,该查询可以解析成显示器112上显示的超过一个节目。例如,对足球节目的请求可能会援引电视机节目类别中列出的足球赛或电影类别中列出的足球电影。可使用在显示器112上显示的内容——包括显示提示——消除用户请求的歧义并确定用户意图。具体地讲,可以使用电影类别在类别界面2440中具有焦点的事实识别期望的特定足球节目,给定电影类别上的焦点,其可能是足球电影。因此,可以在从语音输入确定用户意图时使用如显示器112上所示具有焦点的媒体的类别(或任何其他媒体分组)。还应当理解,用户可以发出与类别相关联的各种其他请求,例如请求显示特定类别的内容(例如,向我显示喜剧电影,向我显示恐怖电影等)。

[0180] 在其他实施例中,用户可以通过各种其他方式援引显示器112上显示的菜单或媒体项,并可基于所显示的内容类似地确定用户意图。应当理解,还可以进一步结合所显示的内容使用与所显示的内容相关联的元数据(例如,电视机节目描述、电影描述等)、模糊匹配技术、同义匹配等,从语音输入确定用户意图。因此可以适应各种形式——包括自然语言请求——的用户请求,并可以根据本文论述的各种实施例确定用户意图。

[0181] 应当理解,在确定用户意图时,可单独或结合在用户设备102上或与遥控器106相关联的显示器上所显示的内容来使用在显示器112上显示的内容。类似地,应当理解,可以在通信地耦接到电视机顶盒104的各种设备的任一种处接收虚拟助理查询,并可使用在显示器112上显示的内容来确定用户意图,不管哪个设备接收查询。可以类似地在显示器112上或另一个显示器上(例如,用户设备102上)显示查询结果。

[0182] 此外,在本文论述的各种实施例的任一个中,虚拟助理系统能够对菜单导航并选择菜单选项,而无需用户专门打开菜单并导航到菜单项。例如,选项菜单可以出现于选择媒

体内容或菜单按钮之后,例如选择图24中的电影选项2444之后。菜单选项可包括播放媒体以及简单播放媒体的替代方式,例如设置提醒以稍晚观看媒体,设置媒体的录制,向收藏列表添加媒体,隐藏媒体不进一步被观看等。在用户观看菜单上方的内容或具有子菜单选项的内容时,用户可以发出本来要求导航到菜单或子菜单加以选择的虚拟助理命令。例如,观看图24的类别界面2440的用户可以发出与电影选项2444相关联的任何菜单命令而不手动打开关联的菜单。例如,用户可能请求向收藏列表添加足球电影,记录晚间新闻,并设置提醒以观看电影B而一直无需导航到与那些媒体选项相关联的菜单或子菜单,其中可能有此类命令。虚拟助理系统因此可以对菜单和子菜单导航,以便代表用户执行命令,无论那些菜单选项是否出现于显示器112上。这样可以简化用户请求并减少实现期望的菜单功能用户必须做出的点击或选择次数。

[0183] 图25示出了用于利用显示器上所显示的媒体内容控制电视机交互和媒体内容的观看历史的示例性过程2500。在方框2502,可以从用户接收语音输入,该语音输入包括与在电视机显示器上所显示的内容相关联的查询。例如,该语音输入可包括关于出现于系统100的显示器112上的角色、演员、电影、电视机节目、体育赛事、队员等的查询(由电视机顶盒104显示)。例如,图19的转录1916包括与显示器112上的视频480中所示的女演员相关联的查询。类似地,图21的转录2122包括与显示器112上的视频480中的角色相关联的查询。语音输入还可包括与出现于显示器112上的菜单或搜索内容相关联的查询,例如选择特定菜单项或得到关于特定搜索结果的信息的查询。例如,显示的菜单内容可包括图23A和图23B中菜单830的媒体选项832。显示的菜单内容可以类似地包括出现于图24的类别界面2440中的电视机选项2442、电影选项2444和/或音乐选项2446。

[0184] 再次参考图25的过程2500,在方框2504处,可基于显示的内容和媒体内容的观看历史来确定查询的用户意图。例如,可基于显示的或最近显示的电视机节目、体育赛事、电影等的场景,确定用户意图。也可基于显示的或最近显示的菜单或搜索内容来确定用户意图。还可以连同与内容相关联的元数据一起分析显示的内容以确定用户意图。例如,可以单独地或结合与所显示的内容相关联的元数据使用参考图19、图21、图23A、图23B所示和所述的内容,确定用户意图。

[0185] 在方框2506处,可基于确定的用户意图来显示查询结果。例如,可以在显示器112上显示类似于图20的助理响应界面2018中的助理响应2020的结果。在另一个实施例中,可以在图22所示的助理响应界面2224中提供文本和可选择媒体作为结果,例如助理文本响应2226和可选择的视频链接2228。在另一个实施例中,显示查询结果可包括显示或播放所选择的媒体内容(例如,经由电视机顶盒104在显示器112上播放所选择的视频)。因此可以通过各种方式,利用所显示的内容和关联的元数据作为语境,从语音输入确定用户意图。

[0186] 在一些实施例中,可以向用户提供虚拟助理查询建议,例如,以通知用户可用的查询,建议用户可以欣赏的内容,教用户如何使用系统,鼓励用户找到附加媒体内容进行消费,等等。在一些实施例中,查询建议可包括可能命令的一般建议(例如,找到喜剧、为我显示电视机指南、搜索动作电影、打开闭合字幕等)。在其他实施例中,查询建议可包括与显示的内容相关的针对性建议(例如,将这个节目添加到观看列表,经由社交媒体共享这个节目,向我显示这个电影的音轨,向我显示这个顾客正在销售的书籍,向我显示该顾客正在插入的电影预告片等)、用户偏好(例如,闭合字幕的使用等)、用户拥有的内容、用户设备上存

储的内容、通知、提示、媒体内容的观看历史(例如,最近显示的菜单项、最近显示的节目场景、演员的最近出现等)等。可以在任何设备上显示建议,包括经由电视机顶盒104在显示器112上,在用户设备102上,或在与遥控器106相关联的显示器上。此外,可基于在特定时间哪些设备在附近和/或与电视机顶盒104在通信来确定建议(例如,从特定时间在房间中看电视的用户的设备建议内容)。在其他实施例中,可基于各种其他语境信息确定建议,包括一天中的时间、大众来源的信息(例如,在给定时间观看的流行节目)、实况节目(例如,实况体育赛事)、媒体内容的观看历史(例如,观看的最后几次演出,最近观看的搜索结果集合,最近观看的媒体选项组等)或各种其他语境信息的任一种。

[0187] 图26示出了示例性建议界面2650,包括基于内容的虚拟助理查询建议2652。在一个实施例中,可以响应于从请求建议的用户接收的输入,在诸如界面2650的界面中提供查询建议。例如,可以从用户设备102或遥控器106接收请求查询建议的输入。在一些实施例中,输入可包括在用户设备102或遥控器106处接收的按钮按压、按钮双击、菜单选择、语音命令(例如,为我显示一些建议,你能为我做些什么,一些选项是什么等)等。例如,在观看与电视机顶盒104相关联的界面时,用户可以双击遥控器106上的物理按钮以请求查询建议,或者可以双击用户设备102上的物理或虚拟按钮,以请求查询建议。

[0188] 可以在移动图像,例如视频480上方,或任何其他背景内容(例如,菜单、静态图像、暂停视频等)上方显示建议界面2650。如本文论述的其他界面那样,建议界面2650可以通过动画方式从显示器112的底部向上滑动,并可以占据最小量的空间,同时充分传送期望的信息,以便限制对背景中视频480的干扰。在其他实施例中,可以在背景内容静止时提供更大的建议界面(例如,暂停的视频、菜单、图像等)。

[0189] 在一些实施例中,可基于显示的媒体内容或媒体内容的观看历史(例如,电影、电视机节目、体育赛事、最近观看的节目、最近观看的菜单、最近观看的电影场景、播放电视剧集的最近场景等)确定虚拟助理查询建议。例如,图26示出了基于内容的建议2652,可基于背景中显示出的所显示视频480确定它,显示器112上现有角色1910、1912和1914。与所显示的内容相关联的元数据(例如,媒体内容的描述性细节)也可以用于确定查询建议。元数据可包括与所显示的内容相关联的各种信息,包括节目标题、角色列表、演员列表、剧集描述、球队名单、球队排名、节目概要、电影细节、情节描述、导演姓名、出品人姓名、演员出现的时间、体育看台、体育比分、流派、剧集分季列表、相关媒体内容或各种其他关联信息。例如,与视频480相关联的元数据可包括角色1910、1912和1914的角色姓名连同饰演那些角色的女演员。元数据还可包括对视频480情节的描述、对前一集或下一集的描述(其中视频480是一系列中的电视剧集)等。

[0190] 图26示出了各种基于内容的建议2652,可基于视频480和与视频480相关联的元数据在建议界面2650中显示。例如,视频480的角色1910可以名为“Blanche”,可以使用该角色名称编写关于角色Blanche或饰演该角色的女演员的信息的查询建议(例如,“谁是饰演Blanche的女演员?”)。可以从与视频480相关联的元数据识别角色1910(例如,角色列表、演员列表、与演员出现相关联的时间等)。在其他实施例中,可以使用面部识别来识别在给定时间出现于显示器112上的女演员和/或角色。可以提供与媒体自身中的角色相关联的各种其他查询建议,例如与角色的人物、概况、与其他角色的关系等相关的查询。

[0191] 在另一个实施例中,显示器112上出现的男演员或女演员可以被识别(例如,基于

元数据和/或面部识别),并可以提供与该男演员或女演员相关联的查询建议。此类查询建议可包括饰演的人物、表演奖、年龄、他们出现的其他媒体、历史、家庭成员、关系或关于男演员或女演员的各种其他详情的任一种。例如,角色1914可以由名为Whitney Davidson的女演员饰演,女演员的姓名Whitney Davidson可以用于编写查询建议以识别其他电影、电视机节目或该女演员Whitney Davidson出现的其他媒体(例如,“Whitney Davidson还在什么其他影视剧中出现过?”)。

[0192] 在其他实施例中,关于节目的细节可用于编写查询建议。剧集概要、情节摘要、剧集列表、剧集标题、系列标题等可用于编写查询建议。例如,可以提供建议以描述电视机节目的上一集中发生了什么(例如,“上一集发生了什么?”),虚拟助理系统可以向该建议提供剧集概要作为响应,该剧集概要来自基于当前在显示器112上显示的剧集(及其关联元数据)识别的先前剧集。在另一个实施例中,可以提供建议以设置针对下一剧集的记录,可以通过系统基于显示器112上显示的当前播放的剧集识别下一集来完成这项任务。在另一个实施例中,可以提供建议以获得关于当前剧集或显示器112上出现的节目的信息,可以使用从元数据获得的节目标题编写查询建议(例如,“这一集‘Their Show’讲的什么?”或“‘Their Show’讲的什么?”)。

[0193] 在另一个实施例中,可以使用与所显示的内容相关联的类别、流派、等级、奖励、描述等编写查询建议。例如,视频480可以对应于描述为具有女性主角角色的喜剧的电视机节目。可以从这一信息编写查询建议以识别具有类似特性的其他节目(例如,“找到有女主角的其他喜剧。”)。在其他实施例中,可基于用户的订阅、可用于回放的内容(例如,电视机顶盒104上的内容、用户设备102上的内容、可用于流传输的内容等)等确定建议。例如,可基于是否有信息或媒体结果来过滤潜在的查询建议。可以排除掉可能不会获得可播放媒体内容或信息性答案的查询建议,和/或可以提供容易获得信息性答案或可播放媒体内容的查询建议(或者在确定提供哪些建议时对其进行更重的加权)。因此可以通过各种方式使用所显示的内容和关联的元数据来确定查询建议。

[0194] 图27示出了用于确认所建议查询的选择的示例性选择界面2754。在一些示例中,用户可以通过讲出查询,利用按钮选择它们,利用光标导航到它们等,来选择显示的查询建议。响应于选择,可以在确认界面,例如,选择界面2754中简要显示选择的建议。在一个示例中,可以动画显示选择的建议2756以从其在建议界面2650中出现的任何地方移动到图27中所示与命令接收确认490相邻的位置(例如,如箭头所示),并可以从显示器隐藏其他未选择的建议。

[0195] 图28A-28B示出了基于选择的查询的示例性虚拟助理应答界面2862。在一些示例中,可以在答案界面,例如答案界面2862中显示对所选查询的带信息答案。在从建议界面2650或选择界面2754切换时,可以显示过渡界面2858,如图28A所示。具体地讲,在下一条内容从显示器112底部向上滚动时,可以向上滚动界面之内先前显示的内容,离开该界面。例如,可以向上滑动或滚动选择的建议2756,直到其在虚拟助理界面的顶边缘处消失,可以从显示器112的底部滑动或滚动助理结果2860,直到其到达图28B中所示的位置。

[0196] 答案界面2862可包括响应于选择的查询建议的(或响应于任何其他查询的)带信息答案和/或媒体结果。例如,响应于选择的查询建议2756,可以确定并提供助理结果2860。具体地讲,响应于对前一集概要的请求,可基于所显示的内容识别前一集,并可以识别且向

用户提供关联的描述或概要。在例示的示例中,助理结果2860可描述对应于显示器112上的视频480的节目的前一集(例如,“在“Their Show”的203集中,Blanche受邀作为嘉宾演讲人参加大学的心理课。Julia和Melissa突然现身,掀起骚动。”)。也可以通过本文所述的其他方式的任一种呈现带信息的答案和媒体结果(例如,可选择视频链接),或者可以通过各种其他方式呈现结果(例如,大声讲出答案,立即播放内容,显示动画,显示图像等)。

[0197] 在另一示例中,可以使用通知或提示确定虚拟助理查询建议。图29示出了媒体内容通知2964(尽管在确定建议时可以考虑任何通知)和既有基于通知的建议2966又有基于内容的建议2652的建议界面2650(其可包括如上文参考图26所述的相同概念的一些)。在一些示例中,可以分析通知的内容以识别相关媒体相关的名称、标题、主题、动作等。在例示的示例中,通知2964包括通知用户可以显示的替代媒体内容的提示——尤其是体育赛事是实况,且比赛内容可能是用户感兴趣的(例如,“比赛还剩五分钟,Zeta队和Alpha队仍然是平局。”)。在一些示例中,可以在显示器112顶部短暂显示通知。可以将通知从显示器112顶部向下滑动(如箭头所示)到图29所示的位置,显示一定量的时间,并向上往回滑动,在显示器112的顶部消失。

[0198] 通知或提示可以通知用户各种信息,例如可用的替代媒体内容(例如,当前显示器112上可能显示的内容的替代内容),可用的实况电视机节目,新下载的媒体内容,最近添加的订阅内容,从朋友接收的建议,接收从另一设备发送的媒体等。也可基于家庭或观看媒体的所识别(例如,基于用户认证,利用账户选择、语音识别、口令等识别)用户对通知进行个性化。在一个示例中,该系统可以中断显示并显示基于可能期望内容的通知,例如,为——基于用户概况、支持球队、优选运动、观看历史等——可能希望该通知内容的用户显示通知2964。例如,可以从体育数据馈送、新闻出口、社交媒体讨论等获得体育赛事比分、比赛状态、剩余时间等,并可以用于识别可能的替代媒体内容,用于通知用户。

[0199] 在其他示例中,可以经由提示或通知提供流行的媒体内容(例如,在很多用户之间),以建议当前观看的内容的替代内容(例如,通知用户流行节目或用户喜欢的流派的节目刚刚开始或通过其他方式可以观看)。在例示的示例中,用户可能关注Zeta队和Alpha队之一或两者(或者可能关注足球或特定运动、联赛等)。该系统可以确定可用的实况内容匹配用户的偏好(例如,另一个频道上的比赛匹配用户的偏好,该比赛剩余时间很少且比分接近)。该系统然后可以决定通过可能期望的内容的通知2964来提示用户。在一些示例中,用户可以选择通知2964(或通知2964之内的链接)以(例如,使用遥控器按钮、光标、口语请求等)切换到建议的内容。

[0200] 可基于通知,通过分析通知内容以识别与相关媒体相关的术语、名称、标题、主题、动作等,来确定虚拟助理查询建议。然后可以使用识别的信息编写适当的虚拟助理查询建议,例如基于通知2964的基于通知的建议2966。例如,可以显示关于实况体育赛事的精彩结尾的通知。如果用户然后请求查询建议,可以显示建议界面2650,包括观看体育赛事,查询球队统计数据或找到与通知相关的内容的查询建议(例如,改为Zeta/Alpha比赛,Zeta队的统计数据如何,还有什么其他足球赛正在进行等)。基于在通知中识别的感兴趣的特定术语,可以类似地确定各种其他查询建议并向用户提供。

[0201] 也可以从用户设备上的内容来确定与媒体内容(例如,用于经由电视机顶盒104消费)相关的虚拟助理查询建议,还可以在用户设备上提供建议。在一些示例中,可以在连接

到或与电视机顶盒104通信的用户设备上识别可播放的设备内容。图30示出了具有界面1360中的示例性图片和视频内容的用户设备102。可以确定用户设备上有什么内容可以回放,或者可能希望什么内容回放。例如,可基于活动应用(例如,照片和视频应用)识别可播放媒体3068,或者可基于在界面1360上显示或未显示的存储的内容,识别可播放媒体3068(例如,在一些示例中可以从活动应用识别内容,或者在其他示例中,在给定时间并不显示)。可播放媒体3068可包括,例如,视频1362、相册1364和照片1366,其每者都可包括可以向电视机顶盒104发送以便显示或回放的个人用户内容。在其他示例中,可以识别用户设备102上存储或显示的任何照片、视频、音乐、比赛界面、应用界面或其他媒体内容并用于确定查询建议。

[0202] 识别了可播放媒体3068之后,可以确定虚拟助理查询建议并向用户提供。图31示出了用户设备102上的示例性TV助理界面3170,其具有基于可播放用户设备内容且基于独立显示器(例如,与电视机顶盒104相关联的显示器112)上显示的视频内容的虚拟助理查询建议。TV助理界面3170可包括专门用于和媒体内容和/或电视机顶盒104交互的虚拟助理界面。用户可以通过例如在观看界面3170时通过双击物理按钮来请求用户设备102上的请求查询建议。可以类似地使用其他输入以指示用于查询建议的请求。如图所示,助理问候3172可以介绍所提供的查询建议(例如,“这里是用于控制您的电视机体验的一些建议。”)。

[0203] 用户设备102上提供的虚拟助理查询建议可包括基于各种源设备的建议以及通用建议。例如,基于设备的建议3174可包括基于用户设备102上存储的内容(包括在用户设备102上显示的内容)的查询建议。基于内容的建议2652可基于在与电视机顶盒104相关联的显示器112上显示的内容。通用建议3176可包括可能不和特定媒体内容或具有媒体内容的特定设备相关联的通用建议。

[0204] 例如,可基于在用户设备102上识别的可播放内容(例如,视频、音乐、照片、比赛界面、应用界面等)确定基于设备的建议3174。在例示的示例中,可以图30中所示的可播放媒体3068确定基于设备的建议3174。例如,给定相册1364被识别为可播放媒体3068,可以使用相册1364的细节编写查询。该系统可以将内容识别为可以在幻灯片中显示的多个照片的相册,然后可以使用相册的标题(在一些情况下)编写查询建议,以显示特定照片相册的幻灯片(例如,“从您的照片显示“毕业相册”的幻灯片。”)。在一些示例中,该建议可包括内容源的指示(例如,“来自您的照片”,“来自Jennifer的电话”,“来自Daniel的平板电脑”等)。该建议还可以使用其他细节来援引特定内容,例如观看来自特定日期的照片的建议(例如,显示您来自6月21日的照片)。在另一示例中,可以将视频1362识别为可播放媒体3068,可以使用视频的标题(或其他标识信息)编写查询建议以播放视频(例如,“显示来自您的视频的“毕业视频”。”)。

[0205] 在其他示例中,可以识别其他连接设备上可用的内容并用于编写虚拟助理查询建议。例如,可以识别来自连接到公共电视机顶盒104的两个用户设备102的每个的内容,并用于编写虚拟助理查询建议。在一些示例中,用户可以选择让哪些内容可以被系统看到进行共享,并可以对系统隐藏其他内容,以免在查询建议中包括它或通过其他方式使其可以回放。

[0206] 例如,可基于与在电视机顶盒104相关联的显示器112上显示的内容来确定图31的界面3170中所示的基于内容的建议2652。在一些示例中,可以通过与上文参考图26所述的

相同方式确定基于内容的建议2652。在例示的示例中,图31中所示的基于内容的建议可基于显示器112上显示的视频480(例如,如图26中那样)。通过这种方式,可基于任意数量的连接设备上显示或可用的内容导出虚拟助理查询建议。除了针对性建议之外,可以预定且提供通用建议3176(例如,为我显示指南,正在进行什么体育项目,三频道正在演什么等等)。

[0207] 图32示出了示例性建议界面2650,其具有基于连接设备的建议3275,连同与电视机顶盒104相关联的显示器112上显示的基于内容的建议2652。在一些示例中,可以通过与上文参考图26所述的相同方式确定基于内容的建议2652。如上所述,可基于任意数量的连接设备上的内容编写虚拟助理查询建议,并可以在任意数量的连接设备上提供建议。图32示出了可以从用户设备102上的内容导出的基于连接设备的建议3275。例如,可以在用户设备102上识别可播放内容,例如界面1360中所示的照片和视频内容,作为图30中的可播放媒体3068。然后可以使用用户设备102上识别的可播放内容编写可以在与电视机顶盒104相关联的显示器112上显示的建议。在一些示例中,可以通过与上文参考图31所述的基于设备的建议3174以相同方式确定基于连接设备的建议3275。此外,如上所述,在一些示例中,可以在建议中包括标识源信息,例如,如基于连接设备的建议3275中所示的“来自Jake的电话”。于是,可基于来自另一个设备的内容(例如,显示的内容、存储的内容等)导出一个设备上提供的虚拟助理查询建议。应当理解,连接设备可包括电视机顶盒104和/或用户设备102可以访问的远程存储设备(例如,访问云中存储的媒体内容以编写建议)。

[0208] 应当理解,可以响应于对建议的请求提供来自各种源的虚拟助理查询建议的任意组合。例如,可以随机组合来自各种源的建议,或者可基于流行性、用户偏好、选择历史等呈现建议。此外,可基于各种其他因素,例如查询历史、用户偏好、查询流行性等,通过各种其他方式确定查询并进行呈现。此外,在一些示例中,可以通过利用延迟之后的新替代建议替换显示的建议来自动循环查询建议。进一步应当理解,用户可以例如通过在触摸屏上轻击、讲出查询、利用导航按键选择查询、利用按钮选择查询、利用光标选择查询等,来选择任何界面上的建议,然后可以提供关联的响应(例如,带信息的和/或媒体响应)。

[0209] 在各示例的任一个中,还可基于可用内容过滤虚拟助理查询建议。例如,会导致不可用媒体内容(例如,没有有线订阅)或可能没有关联的带信息答案的潜在查询建议可以被取消建议资格并保持不被显示。另一方面,会导致用户可以访问的可立即播放媒体内容的潜在查询建议,相对于其他潜在建议,可以被加权或通过其他方式偏置以进行显示。通过这种方式,也可以在确定用于显示的虚拟助理查询建议时使用供用户观看的媒体内容的可用性。

[0210] 此外,在各示例的任一个中,可以提供预先加载的查询答案以替代或补充建议(例如,在建议界面2650中)。可基于个人使用和/或当前语境选择并提供此类预先加载的查询答案。例如,观看特定节目的用户可以轻击按钮、双击按钮、长按按钮等,以接收建议。作为查询建议的替代或补充,可以自动提供基于语境的信息,例如标识播放歌曲或音轨(例如,“这首歌是表演秀”),标识当前播放剧集的剧组成员(例如,“女演员Janet Quinn饰演Genevieve”),标识类似媒体(例如,“节目Q类似于这个节目”),或者提供本文论述的其他查询的任何结果。

[0211] 此外,可以在各种界面的任一种中提供示能表示,供用户对媒体内容评级,以通知虚拟助理用户的偏好(例如,可选择的评级尺度)。在其他示例中,用户可以讲出评级信息作

为自然语言命令(例如,“我喜欢”,“我不喜欢”,“我不喜欢这个节目”等)。在其他示例中,在本文所示所述的各种界面的任一种中,可以提供各种其他功能和信息元素。例如,界面还可包括通往重要功能和地方的链接,例如搜索链接、购物链接、媒体链接等。在另一示例中,界面还可包括基于当前播放的内容接下来还可以看什么的推荐(例如,选择类似内容)。在又一示例中,界面还可包括基于个性化品味和/或最近活动的接下来还可以看什么的推荐(例如,基于用户评级、用户输入的偏好、最近观看的节目等选择内容)。在其他示例中,界面还可包括用于用户交互的指令(例如,“按下并保持以向虚拟助理讲话,”“轻击一次以得到建议,”等)。在一些示例中,提供预先加载的答案、建议等可以提供令人愉悦的用户体验,同时还使得内容容易被很多用户使用(例如,各种技能水平的用户,不论其语言或其他控制载体为何)。

[0212] 图33示出了用于建议虚拟助理交互以用于控制媒体内容(例如,虚拟助理查询)的示例性过程3300。在方框3302,可以在显示器上显示媒体内容。例如,如图26所示,可以经由电视机顶盒104在显示器112上显示视频480,或者可以在用户设备102的触摸屏246上显示界面1360,如图30中所示。在方框3304,可以从用户接收输入。该输入可包括对虚拟助理查询建议的请求。该输入可包括按钮按压、按钮双击、菜单选择、对建议的口语查询等。

[0213] 在方框3306,可基于媒体内容和/或媒体内容的观看历史来确定虚拟助理查询。例如,可基于显示的节目、菜单、应用、媒体内容列表、通知等确定虚拟助理查询。在一个示例中,可基于视频480和参考图26所述的关联元数据来确定基于内容的建议2652。在另一示例中,可基于参考图29所述的通知2964确定基于通知的建议2966。在又一示例中,可基于上文参考图30和31所述的用户设备102上的可播放媒体3068确定基于设备的建议3174。在其他示例中,可基于上文参考图32所述的用户设备102上的可播放媒体3068确定基于连接设备的建议3275。

[0214] 再次参考图33的过程3300,在方框3308处,可以在显示器上显示虚拟助理查询。例如,可以如参考图26、27、29、31和32所示和所述显示所确定的查询建议。如上所述,可基于各种其他信息确定和显示查询建议。此外,可基于来自具有另一显示器的另一设备的内容导出显示器上提供的虚拟助理查询建议。于是可以向用户提供针对性的虚拟助理查询建议,由此辅助用户获知潜在查询以及提供期望的内容建议等益处。

[0215] 此外,在本文所述的各个示例的任一个中,可以针对特定用户对各个方面进行个性化。可以使用用户数据,包括联系人、偏好、位置、收藏媒体等,解释语音命令并方便用户与本文论述的各种设备交互。也可以根据用户偏好、联系人、文本、使用历史、概况数据、人口统计信息等,通过各种方式修改本文论述的各种过程。此外,可以随着时间,基于用户交互(例如,频繁讲出的命令、频繁选择的应用等)更新这样的偏好和设置。收集和使用可从各种源得到的用户数据可以用于改善向用户传递邀请内容或他们可能感兴趣的任何其他内容。本公开预期,在一些实例中,这些所采集的数据可包括唯一地识别或可用于联系或定位特定人员的个人信息数据。此类个人信息数据可包括人口数据、基于位置的数据、电话号码、电子邮件地址、家庭地址或任何其他识别信息。

[0216] 本公开认识到在本发明技术中使用此类个人信息数据可用于使用户受益。例如,该个人信息数据可用于递送用户较感兴趣的目标内容。因此,使用此类个人信息数据使得能够对所递送的内容进行有计划的控制。此外,本公开还预期个人信息数据有益于用户的

其他用途。

[0217] 本公开还预期负责此类个人信息数据的收集、分析、公开、传输、存储或其他用途的实体将遵守已确立的隐私政策和/或隐私实践。具体地讲，此类实体应当实行并坚持使用被公认为满足或超出对维护个人信息数据的隐私性和安全性的行业或政府要求的隐私政策和实践。例如，来自用户的个人信息应当被收集用于实体的合法且合理的用途，并且不在这些合法使用之外共享或出售。另外，此类收集应当仅在用户知情同意之后进行。另外，此类实体应采取任何所需的步骤，以保障和保护对此类个人信息数据的访问，并且确保能够访问个人信息数据的其他人遵守他们的隐私政策和程序。另外，这种实体可使其本身经受第三方评估以证明其遵守广泛接受的隐私政策和实践。

[0218] 不管前述情况如何，本公开还预期用户选择性地阻止使用或访问个人信息数据的示例。即本公开预期可提供硬件元件和/或软件元件，以防止或阻止对此类个人信息数据的访问。例如，就广告递送服务而言，本发明的技术可被配置为在注册服务期间允许用户选择“加入”或“退出”参与对个人信息数据的收集。又如，用户可选择不为目标内容递送服务提供位置信息。再如，用户可选择不提供精确的位置信息，但准许传输位置区域信息。

[0219] 因此，虽然本公开广泛地覆盖了使用个人信息数据来实现一个或多个各种所公开的实施例，但本公开还预期各种实施例也可在无需访问此类个人信息数据的情况下被实现。即，本发明技术的各种实施例不会由于缺少此类个人信息数据的全部或一部分而无法正常进行。例如，可通过基于非个人信息数据或绝对最低限度的个人信息诸如与用户相关联的设备所请求的内容、对内容递送服务可用的其他非个人信息或公开可用信息来推断偏好，从而选择内容并递送给用户。

[0220] 根据一些示例，图34示出了电子设备3400的功能框图，根据各种所述示例的原理配置该电子设备，例如，以使用虚拟助理来控制电视机交互并利用不同的界面显示关联的信息。设备的功能块可由执行各种所述实施例的原理的硬件、软件或硬件和软件的组合来实现。本领域的技术人员应当理解，图34中所述的功能块可被组合或被分离为子块，以便实现各种所述实施例的原理。因此，本文的具体实施方式任选地支持本文所述的功能块的任何可能的组合或分离或进一步限定。

[0221] 如图34所示，电子设备3400可包括配置为显示媒体、界面和其他内容的显示单元3402（例如，显示器112、触摸屏246等）。电子设备3400还可包括配置为接收信息的输入单元3404（例如，麦克风、接收机、触摸屏、按钮等），所述信息例如是语音输入、触觉输入、手势输入等。电子设备3400还可包括耦接至显示单元3402和输入单元3404的处理单元3406。在一些示例中，处理单元3406可包括语音输入接收单元3408、媒体内容确定单元3410、第一用户界面显示单元3412、选择接收单元3414和第二用户界面显示单元3416。

[0222] 处理单元3406可被配置为从用户（例如，经由输入单元3404）接收语音输入。处理单元3406还可被配置为（例如，使用媒体内容确定单元3410）基于语音输入来确定媒体内容。处理单元3406还可被配置为（例如，使用第一用户界面显示单元3412在显示单元3402上）显示具有第一尺寸的第一用户界面，其中第一用户界面包括一个或多个至媒体内容的可选择链接。处理单元3406还可被配置为（例如，使用选择接收单元3414从输入单元3404）接收对一个或多个可选择链接的选择。处理单元3406还可被配置为，响应于该选择，（例如，使用第二用户界面显示单元3416在显示单元3402上）显示具有比第一尺寸大的第二尺寸的

第二用户界面，其中第二用户界面包括与该选择相关联的媒体内容。

[0223] 在一些示例中，(例如，第一用户界面显示单元3412的)第一用户界面响应于(例如，选择接收单元3414的)该选择而扩展到(例如，第二用户界面显示单元3416的)第二用户界面中。在其他示例中，第一用户界面重叠在播放媒体内容上。在一个示例中，第二用户界面重叠在播放媒体内容上。在另一示例中，(例如，语音输入接收单元3408来自输入单元3404的)语音输入包括查询，并且(例如，媒体内容确定单元3410的)媒体内容包括查询结果。在又一示例中，第一用户界面包括除至媒体内容的一个或多个可选择链接之外的至查询结果的链接。在其他示例中，该查询包括关于天气的查询，并且第一用户界面包括至与关于天气的查询相关联的媒体内容的链接。在另一示例中，该查询包括位置，并且至与关于天气的查询相关联的媒体内容的链接包括至与该位置处的天气相关联的媒体内容的一部分的链接。

[0224] 在一些示例中，响应于该选择，处理单元3406可被配置为播放与该选择相关联的媒体内容。在一个示例中，媒体内容包括电影。在另一示例中，媒体内容包括电视机节目。在另一示例中，媒体内容包括体育赛事。在一些示例中，(例如，第二用户界面显示单元3416的)第二用户界面包括对与该选择相关联的媒体内容的描述。在其他示例中，第一用户界面包括至购买媒体内容的链接。

[0225] 处理单元3406还可被配置为从用户(例如，经由输入单元3404)接收附加语音输入，其中附加语音输入包括与所显示的内容相关联的查询。处理单元3406还可被配置为基于与所显示的内容相关联的元数据来确定对与所显示的内容相关联的查询的响应。处理单元3406还可被配置为响应于接收到附加语音输入，(例如，在显示单元3402上)显示第三用户界面，其中第三用户界面包括对与所显示的内容相关联的查询的所确定的响应。

[0226] 处理单元3406还可被配置为接收用于发起(例如，经由输入单元3404)对语音输入的接收的指示。处理单元3406还可被配置为响应于接收到该指示，(例如，在显示单元3402上)显示准备就绪确认。处理单元3406还可被配置为响应于接收到语音输入而显示侦听确认。处理单元3406还可被配置为检测语音输入的结束，响应于检测到语音输入结束，显示处理确认。在一些示例中，处理单元3406还可被配置为显示对语音输入的转录。

[0227] 在一些示例中，电子设备3400包括电视机。在其他示例中，电子设备3400包括电视机顶盒。在其他示例中，电子设备3400包括遥控器。在其他示例中，电子设备3400包括移动电话。

[0228] 在一个示例中，(例如，第一用户界面显示单元3412的)第一用户界面中的一个或多个可选择链接包括与媒体内容相关联的移动图像。在一些示例中，与媒体内容相关联的移动图像包括对媒体内容的实况馈送。在其他示例中，第一用户界面中的一个或多个可选择链接包括与媒体内容相关联的静态图像。

[0229] 在一些示例中，处理单元3406还可被配置为确定当前所显示的内容是否包括移动图像或控制菜单；响应于确定当前所显示的内容包括移动图像，选择小尺寸作为(例如，第一用户界面显示单元3412的)第一用户界面的第一尺寸；响应于确定当前所显示的内容包括控制菜单，选择比小尺寸更大的大尺寸作为(例如，第一用户界面显示单元3412的)第一用户界面的第一尺寸。在其他示例中，处理单元3406还可被配置为基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示，并显示包括所

确定的替代媒体内容的通知。

[0230] 根据一些示例,图35示出了电子设备3500的功能框图,根据各种所述示例的原理配置该电子设备,例如,以使用虚拟助理和多个用户设备来控制电视机交互。设备的功能块可由执行各种所述实施例的原理的硬件、软件或硬件和软件的组合来实现。本领域的技术人员应当理解,图35中所述的功能块可被组合或被分离为子块,以便实现各种所述实施例的原理。因此,本文的具体实施方式任选地支持本文所述的功能块的任何可能的组合或分离或进一步限定。

[0231] 如图35所示,电子设备3500可包括配置为显示媒体、界面和其他内容的显示单元3502(例如,显示器112、触摸屏246等)。电子设备3500还可包括配置为接收信息的输入单元3504(例如,麦克风、接收机、触摸屏、按钮等),所述信息例如是语音输入、触觉输入、手势输入等。电子设备3500还可包括耦接至显示单元3502和输入单元3504的处理单元3506。在一些示例中,处理单元3506可包括语音输入接收单元3508、用户意图确定单元3510、媒体内容确定单元3512和媒体内容播放单元3514。

[0232] 处理单元3506可被配置为(例如,使用语音输入接收单元3508从输入单元3504)从具有第一显示器(例如,在一些示例中为显示单元3502)的第一设备(例如,设备3500)处的用户接收语音输入。处理单元3506还可被配置为基于在第一显示器上所显示的内容(例如,利用用户意图确定单元3510)确定语音输入的用户意图。处理单元3506还可被配置为(例如,使用媒体内容确定单元3512)基于用户意图来确定媒体内容。处理单元3506还可被配置为(例如,利用媒体内容播放单元3514)在与第二显示器(例如,一些示例中的显示单元3502)相关联的第二设备上播放媒体内容。

[0233] 在一个示例中,第一设备包括遥控器。在另一示例中,第一设备包括移动电话。在另一示例中,第一设备包括平板电脑。在一些示例中,第二设备包括电视机顶盒。在其他示例中,第二显示器包括电视机。

[0234] 在一些示例中,在第一显示器上所显示的内容包括应用界面。在一个示例中,(例如,来自输入单元3504的语音输入接收接收单元3508)语音输入包括用于播放与应用界面相关联的媒体的请求。在一个示例中,媒体内容包括与应用界面相关联的媒体。在另一示例中,该应用界面包括相册,并且媒体包括相册中的一个或多个照片。在又一示例中,该应用界面包括一个或多个视频的列表,并且媒体包括一个或多个视频中的一个视频。在其他示例中,该应用界面包括电视机节目列表,并且媒体包括电视机节目列表中的电视机节目。

[0235] 在一些示例中,处理单元3506还可被配置为确定第一设备是否被授权;其中响应于确定第一设备被授权而在第二设备上播放媒体内容。处理单元3506还可被配置为基于语音输入来识别用户,并基于与所识别用户相关联的数据(例如,利用用户意图确定单元3510)确定语音输入的用户意图。处理单元3506还可被配置为基于语音输入来确定用户是否被授权;其中响应于确定用户是授权用户而在第二设备上播放媒体内容。在一个示例中,确定用户是否被授权包括使用语音识别来分析语音输入。

[0236] 在其他示例中,处理单元3506还可被配置为响应于确定用户意图包括对信息的请求,在第一设备的第一显示器上显示与媒体内容相关联的信息。处理单元3506还可被配置为响应于确定用户意图包括用于播放媒体内容的请求,在第二设备上播放媒体内容。

[0237] 在一些示例中,语音输入包括用于在第二设备上播放内容的请求,响应于用于在

第二设备上播放内容的请求,在第二设备上播放媒体内容。处理单元3506还可被配置为基于媒体格式、用户偏好或默认设置来确定是否应当在第一显示器或第二显示器上显示所确定的媒体内容。在一些示例中,响应于确定应当在第二显示器上显示所确定的媒体内容,在第二显示器上显示媒体内容。在其他示例中,响应于确定应当在第一显示器上显示所确定的媒体内容,在第一显示器上显示媒体内容。

[0238] 在其他示例中,处理单元3506还可被配置为确定包括第二设备和第三设备的两个或更多个设备中的每个设备的邻近性。在一些示例中,基于相对于第三设备邻近性的第二设备的邻近性,在与第二显示器相关联的第二设备上播放媒体内容。在一些示例中,确定两个或更多个设备中的每个设备的邻近性包括基于蓝牙LE来确定邻近性。

[0239] 在一些示例中,处理单元3506还可被配置为显示包括与第二显示器相关联的第二设备的显示设备的列表,并接收显示设备的列表中对第二设备的选择。在一个示例中,响应于接收到对第二设备的选择在第二显示器上显示媒体内容。处理单元3506还可被配置为确定耳机是否附接到第一设备。处理单元3506还可被配置为响应于确定耳机附接到第一设备,在第一显示器上显示媒体内容。处理单元3506还可被配置为响应于确定耳机未附接到第一设备,在第二显示器上显示媒体内容。在其他示例中,处理单元3506还可被配置为基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示,并显示包括所确定的替代媒体内容的通知。

[0240] 根据一些示例,图36示出了电子设备3600的功能框图,根据各种所述示例的原理配置该电子设备,例如,以利用显示器上示出的媒体内容和媒体内容的观看历史控制电视机交互。设备的功能块可由执行各种所述实施例的原理的硬件、软件或硬件和软件的组合来实现。本领域的技术人员应当理解,图36中所述的功能块可被组合或被分离为子块,以便实现各种所述实施例的原理。因此,本文的具体实施方式任选地支持本文所述的功能块的任何可能的组合或分离或进一步限定。

[0241] 如图36所示,电子设备3600可包括配置为显示媒体、界面和其他内容的显示单元3602(例如,显示器112、触摸屏246等)。电子设备3600还可包括配置为接收信息的输入单元3604(例如,麦克风、接收机、触摸屏、按钮等),所述信息例如是语音输入、触觉输入、手势输入等。电子设备3600还可包括耦接至显示单元3602和输入单元3604的处理单元3606。在一些示例中,处理单元3606可包括语音输入接收单元3608、用户意图确定单元3610和查询结果显示单元3612。

[0242] 处理单元3606可被配置为(例如,利用语音输入接收单元3608从输入单元3604)从用户接收语音输入,其中该语音输入包括与在电视机显示器(例如,一些示例中的显示单元3602)上显示的内容相关联的查询。处理单元3606还可被配置为基于在电视机显示器上所显示的内容和媒体内容的观看历史中的一个或多个(例如,利用用户意图确定单元3610)确定查询的用户意图。处理单元3606还可以配置为基于所确定的用户意图来显示(例如,使用查询结果显示单元3612)查询结果。

[0243] 在一个示例中,在遥控器处接收语音输入。在另一示例中,在移动电话处接收语音输入。在一些示例中,在电视机显示屏上显示查询结果。在另一示例中,在电视机显示器上示出的内容包括电影。在又一示例中,在电视机显示器上示出的内容包括电视机节目。在一示例中,在电视机显示器上示出的内容包括体育赛事。

[0244] 在一些示例中,该查询包括对关于与在电视机显示器上所显示的内容相关联的人的信息的请求,并且(例如,查询结果显示单元3612的)查询结果包括关于该人的信息。在一个示例中,查询结果包括与该人相关联的媒体内容。在另一示例中,媒体内容包括与该人相关联的电影、电视机节目或体育赛事中的一者或多者。在一些示例中,该查询包括对关于在电视机显示器上所显示的内容中的角色的信息的请求,并且查询结果包括关于该角色的信息或关于饰演该角色的演员的信息。在一个示例中,查询结果包括与饰演该角色的演员相关联的媒体内容。在另一示例中,媒体内容包括与饰演该角色的演员相关联的电影、电视机节目或体育赛事中的一者或多者。

[0245] 在一些示例中,处理单元3606还可被配置为基于与在电视机显示器上所显示的内容相关联的元数据或媒体内容的观看历史来确定查询结果。在一个示例中,元数据包括标题、描述、角色列表、演员列表、队员列表、流派、或与在电视机显示器上所显示的内容相关联的显示调度或媒体内容的观看历史。在另一示例中,在电视机显示器上所显示的内容包括媒体内容列表,并且查询包括用于显示列表中的项目中的一个项目的请求。在又一示例中,在电视机显示器上所显示的内容还包括媒体内容列表中的具有焦点的项目,并且(例如,使用用户意图确定单元3610)确定查询的用户意图包括识别具有焦点的项目。在一些示例中,处理单元3606还可被配置为基于电视机显示器上最近显示的菜单或搜索内容(例如,利用用户意图确定单元3610)确定查询的用户意图。在一个示例中,在电视机显示器上所显示的内容包括一页列出的媒体,并且最近显示的菜单或搜索内容包括前一页列出的媒体。在另一示例中,在电视机显示器上所显示的内容包括一种或多种类别的媒体,并且该一种或多种类别的媒体中的一种类别的媒体具有焦点。在一个示例中,处理单元3606还可被配置为基于该一种或多种类别的媒体中具有焦点的一种(例如,利用用户意图确定单元3610)确定查询的用户意图。在另一示例中,媒体的类别包括电影、电视机节目和音乐。在其他示例中,处理单元3606还可被配置为基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示,并显示包括所确定的替代媒体内容的通知。

[0246] 根据一些示例,图37示出了电子设备3700的功能框图,根据各种所述示例的原理配置该电子设备,例如,以建议用于控制媒体内容的虚拟助理交互。设备的功能块可由执行各种所述实施例的原理的硬件、软件或硬件和软件的组合来实现。本领域的技术人员应当理解,图37中所述的功能块可被组合或被分离为子块,以便实现各种所述实施例的原理。因此,本文的具体实施方式任选地支持本文所述的功能块的任何可能的组合或分离或进一步限定。

[0247] 如图37所示,电子设备3700可包括配置为显示媒体、界面和其他内容的显示单元3702(例如,显示器112、触摸屏246等)。电子设备3700还可包括配置为接收信息的输入单元3704(例如,麦克风、接收机、触摸屏、按钮等),所述信息例如是语音输入、触觉输入、手势输入等。电子设备3700还可包括耦接至显示单元3702和输入单元3704的处理单元3706。在一些示例中,处理单元3706可包括媒体内容显示单元3708、输入接收单元3710、查询确定单元3712和查询显示单元3714。

[0248] 处理单元3706可被配置为在显示器(例如,显示单元3702)上(例如,使用媒体内容显示单元3708)显示媒体内容。处理单元3706还可被配置为(例如,使用输入接收单元3710从输入单元3704)从用户接收输入。处理单元3706还可被配置为基于媒体内容和媒体内容

的观看历史中的一个或多个(例如,利用查询确定单元3712)确定一个或多个虚拟助理查询。处理单元3706还可以配置为(例如,使用查询显示单元3714)在显示器上显示一个或多个虚拟助理查询。

[0249] 在一个示例中,在遥控器上从用户接收输入。在另一示例中,在移动电话上从用户接收输入。在一些示例中,一个或多个虚拟助理查询重叠在移动图像上。在另一示例中,输入包括对按钮的双击。在一个示例中,媒体内容包括电影。在另一示例中,媒体内容包括电视机节目。在又一示例中,媒体内容包括体育赛事。

[0250] 在一些示例中,一个或多个虚拟助理查询包括关于媒体内容中出现的人的查询。在其他示例中,一个或多个虚拟助理查询包括关于媒体内容中出现的角色的查询。在另一示例中,一个或多个虚拟助理查询包括对与媒体内容中出现的人相关联的媒体内容的查询。在一些示例中,媒体内容或媒体内容的观看历史包括电视机节目的剧集,并且一个或多个虚拟助理查询包括关于电视机节目的另一剧集的查询。在另一示例中,媒体内容或媒体内容的观看历史包括电视机节目的剧集,并且一个或多个虚拟助理查询包括用于设置提醒以观看或录制媒体内容的后续剧集的请求。在又一示例中,一个或多个虚拟助理查询包括对媒体内容的描述性详情的查询。在一个示例中,描述性详情包括节目标题、角色列表、演员列表、剧集描述、团队名单、团队评级或节目概要中的一者或者多者。

[0251] 在一些示例中,处理单元3706还可以配置为接收对一个或多个虚拟助理查询中的一个虚拟助理查询的选择。处理单元3706还可被配置为显示一个或多个虚拟助理查询中的所选择的一个虚拟助理查询的结果。在一个示例中,确定一个或多个虚拟助理查询包括基于以下各项中的一者或者多者来确定一个或多个虚拟助理查询:查询历史、用户偏好或查询流行性。在另一个示例中,确定一个或多个虚拟助理查询包括基于用户可观看的媒体内容来确定一个或多个虚拟助理查询。在又一个示例中,确定一个或多个虚拟助理查询包括基于所接收的通知来确定一个或多个虚拟助理查询。在又一个示例中,确定一个或多个虚拟助理查询包括基于活动应用程序来确定一个或多个虚拟助理查询。在其他示例中,处理单元3706还可被配置为基于用户偏好、节目流行性和实况体育赛事的状态中的一种或多种来确定替代媒体内容以用于显示,并显示包括所确定的替代媒体内容的通知。

[0252] 尽管已经参考附图充分描述了示例,但要指出的是,各种改变和修改对于本领域的技术人员而言将变得显而易见(例如,修改根据结合本文论述的任何其他系统或过程描述的概念在本文论述的任何系统或过程)。应当理解,此类变化和修改被认为包括在由所附权利要求所限定的各种示例的范围内。

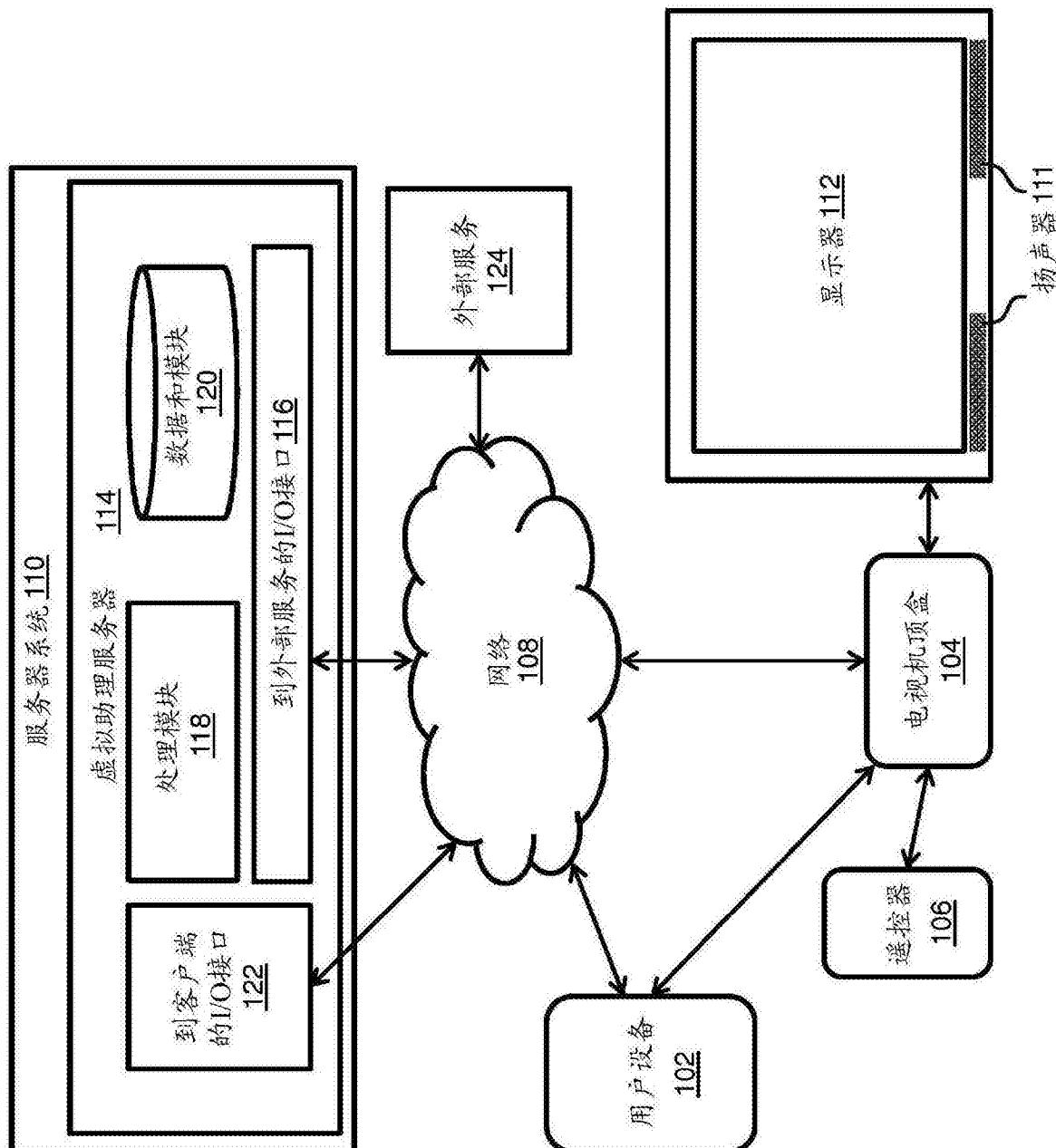


图 1

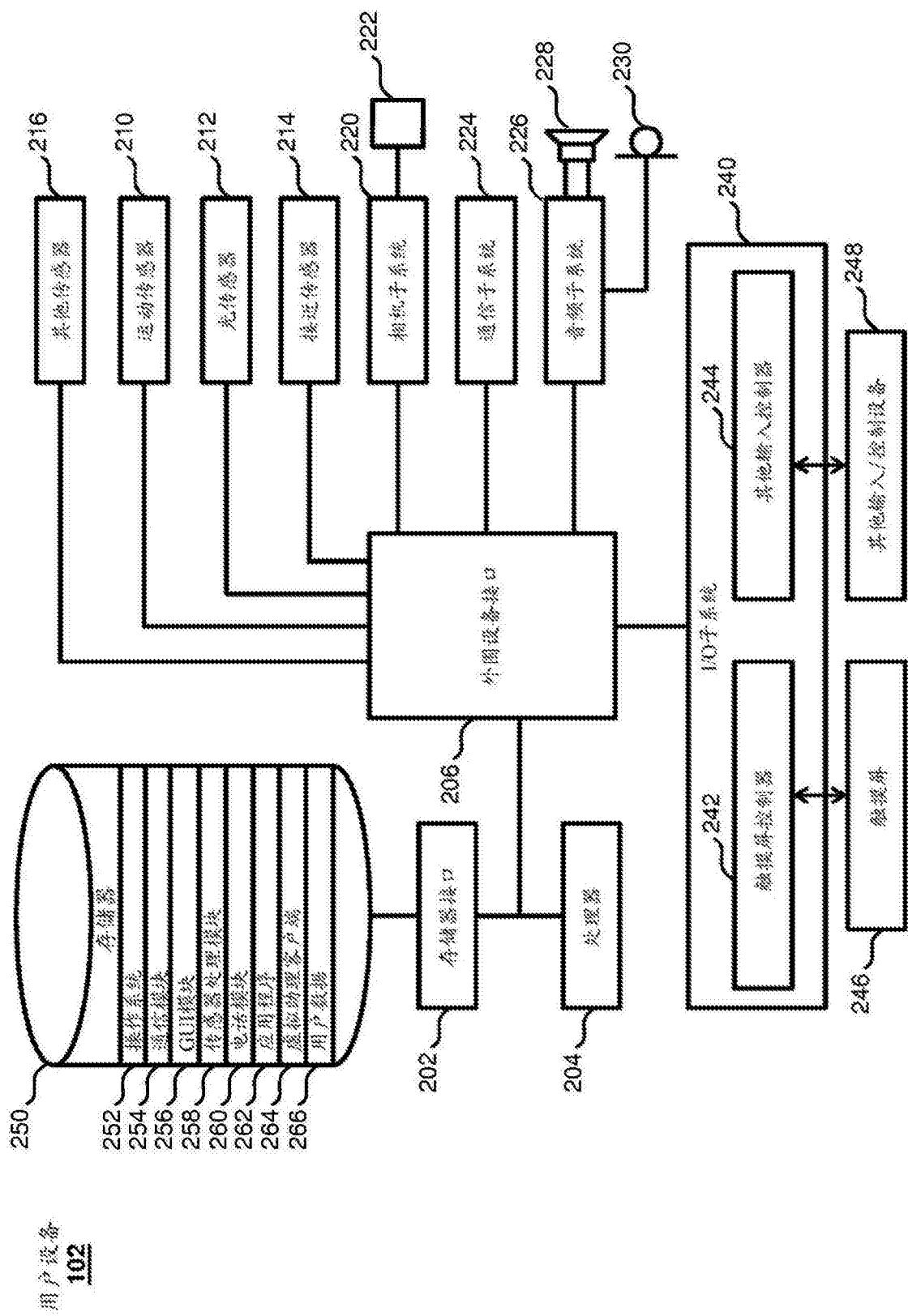


图2

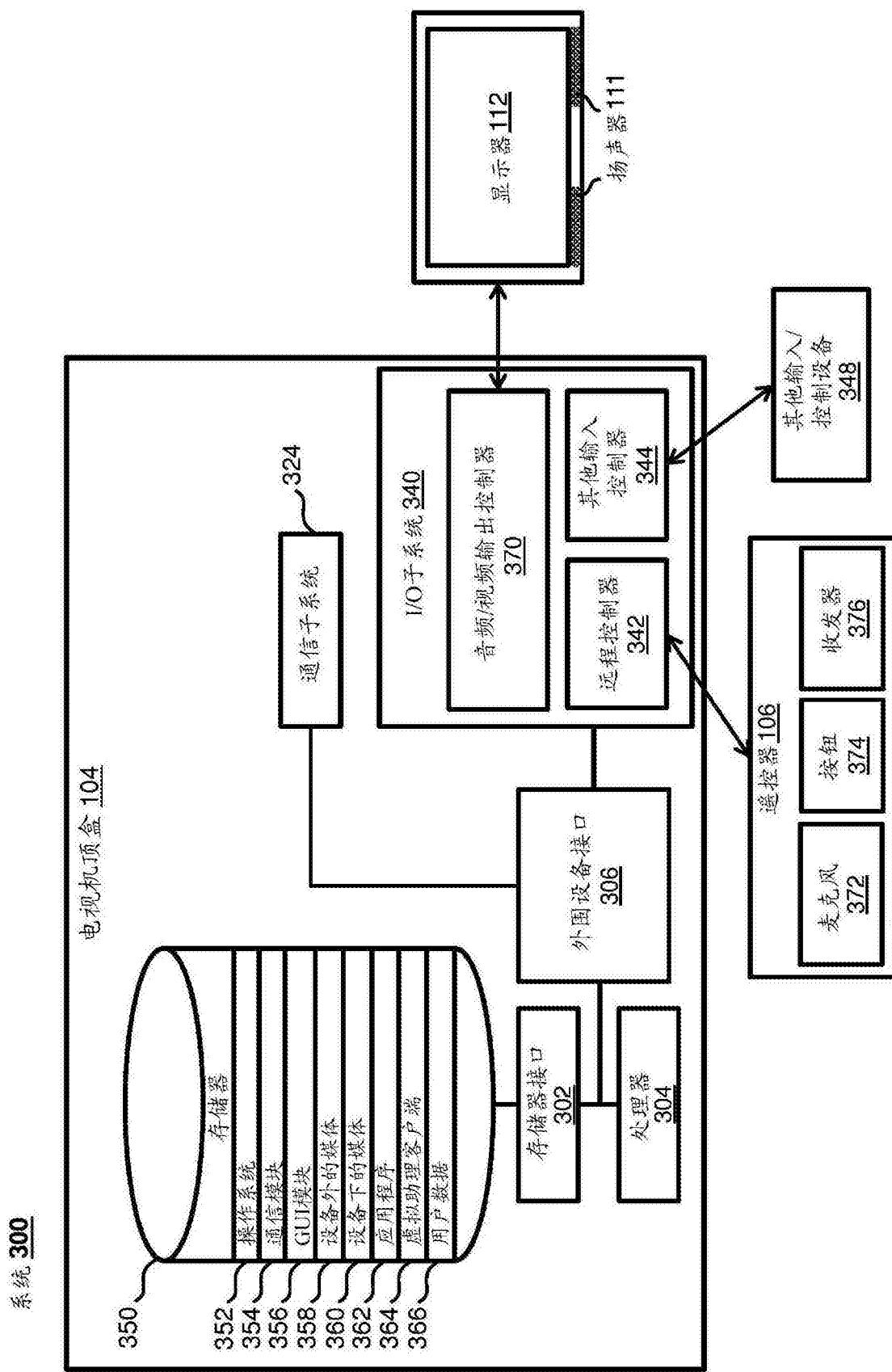


图3

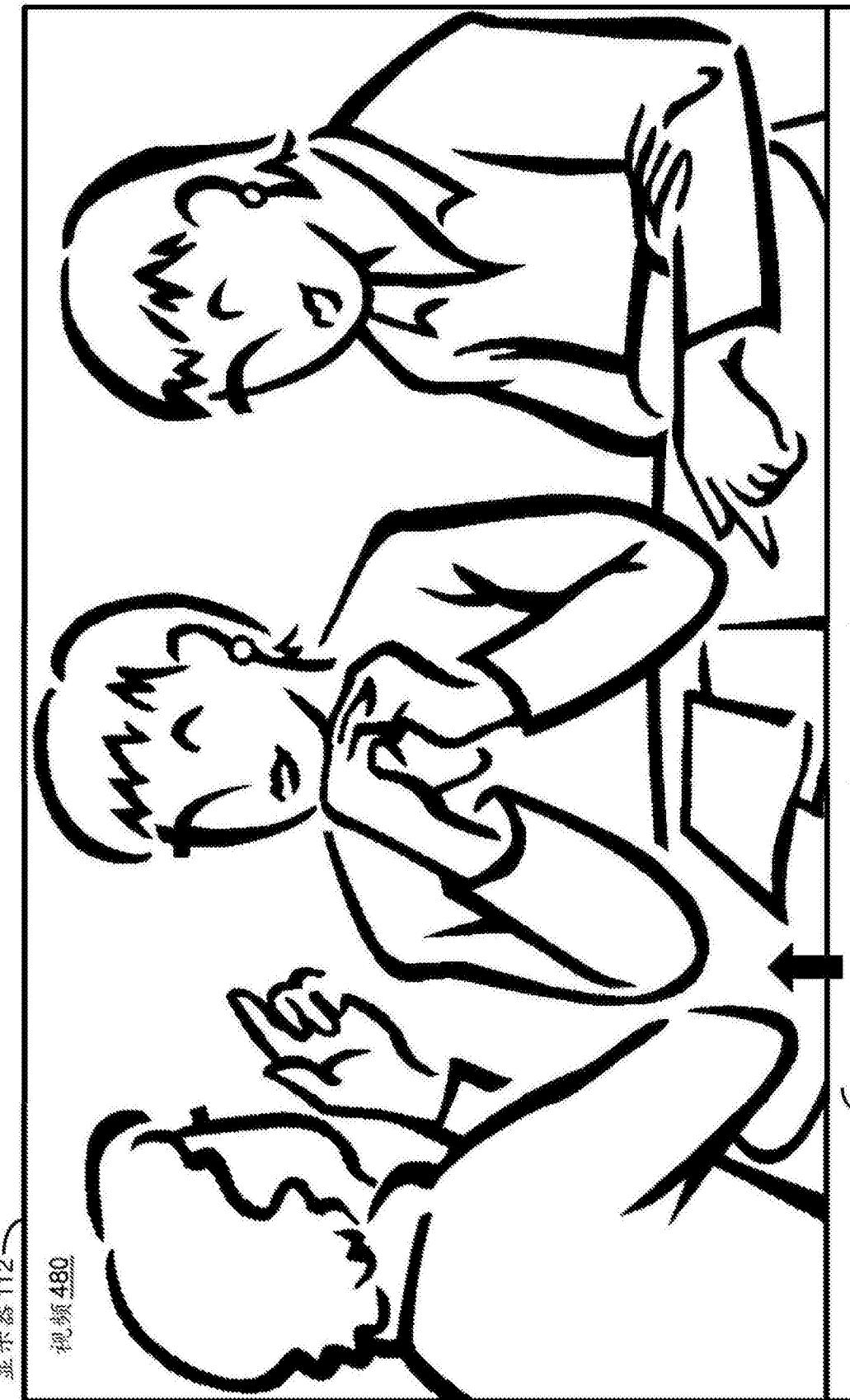


图4A

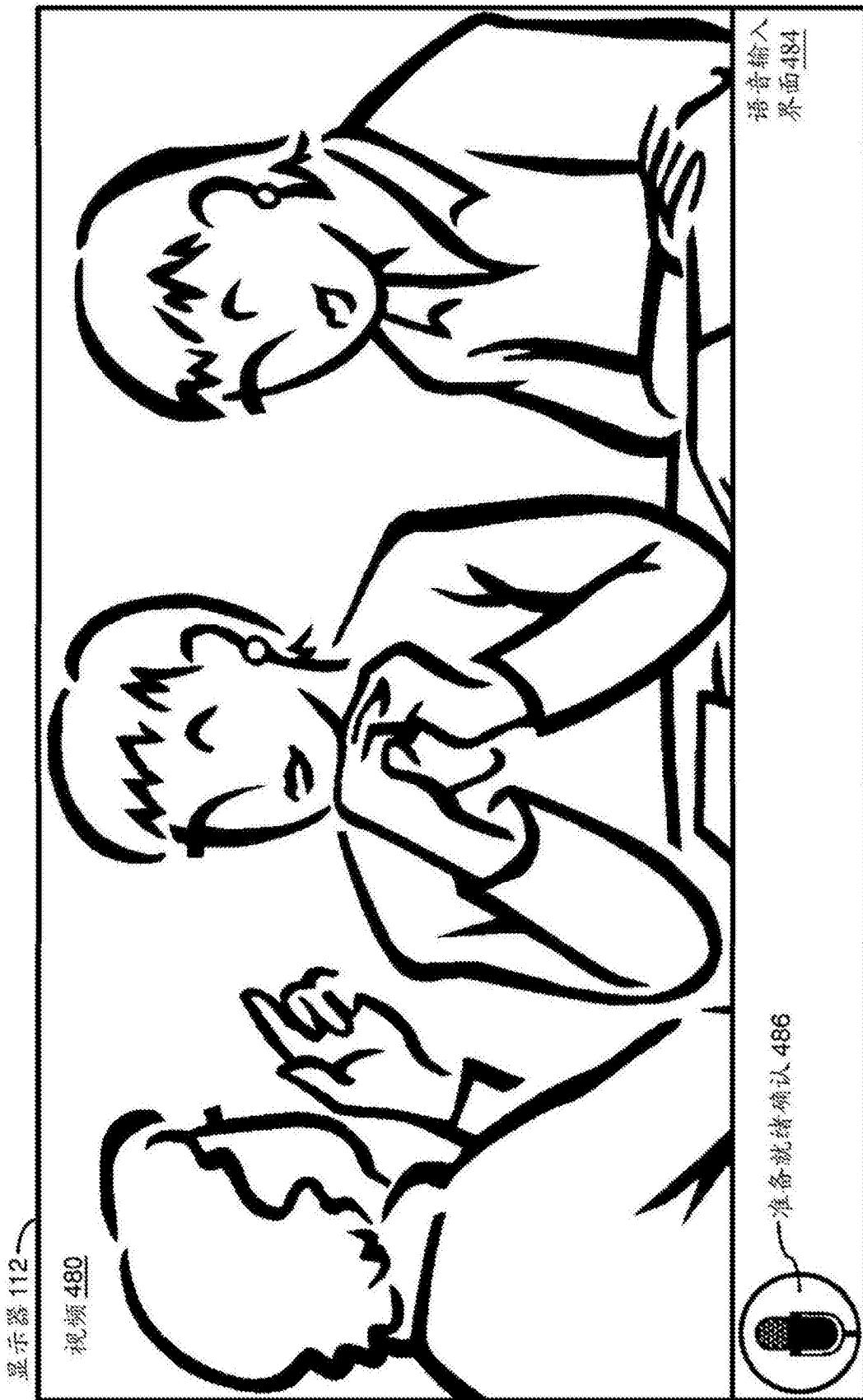


图4B



图4C

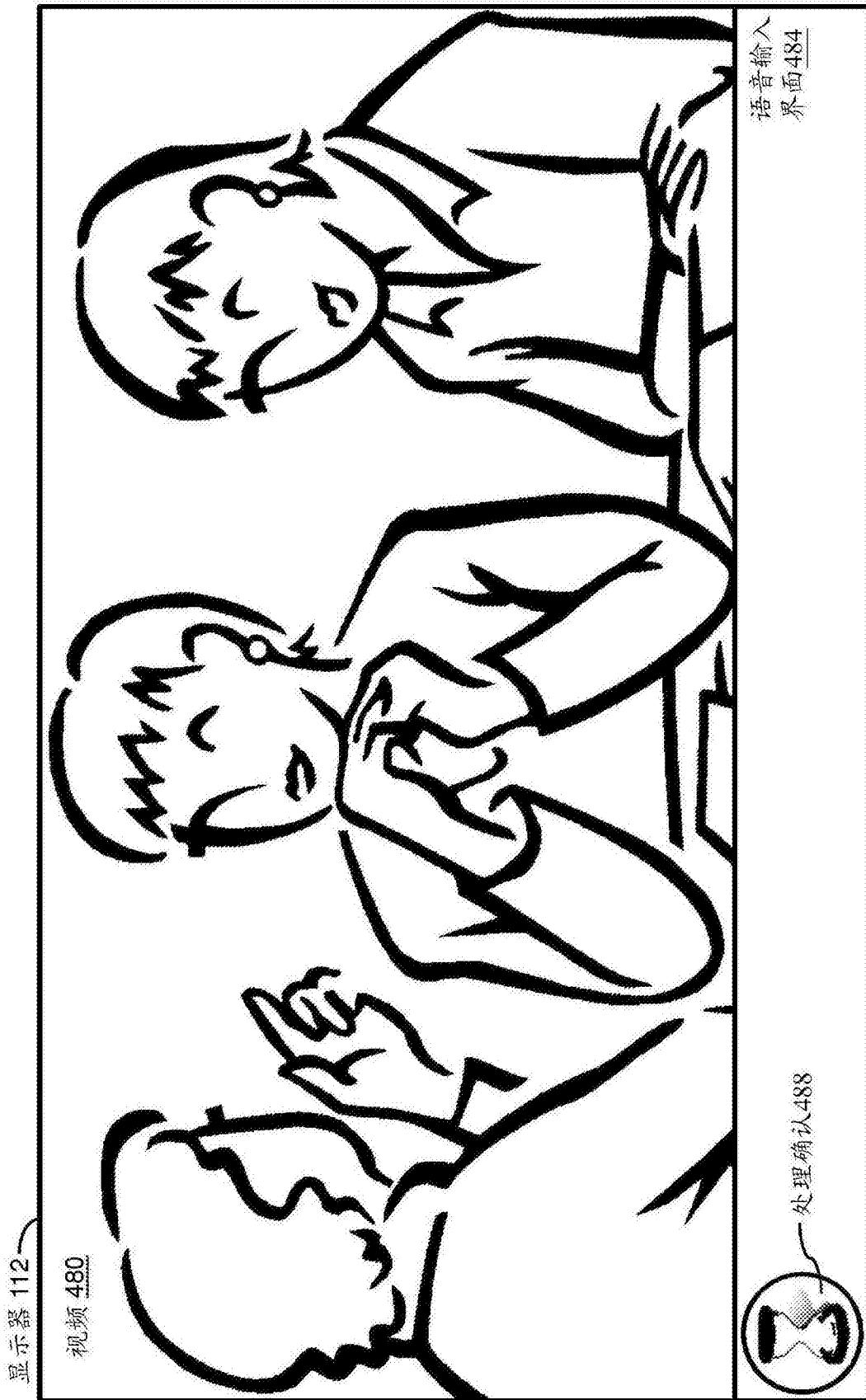


图4D



图4E

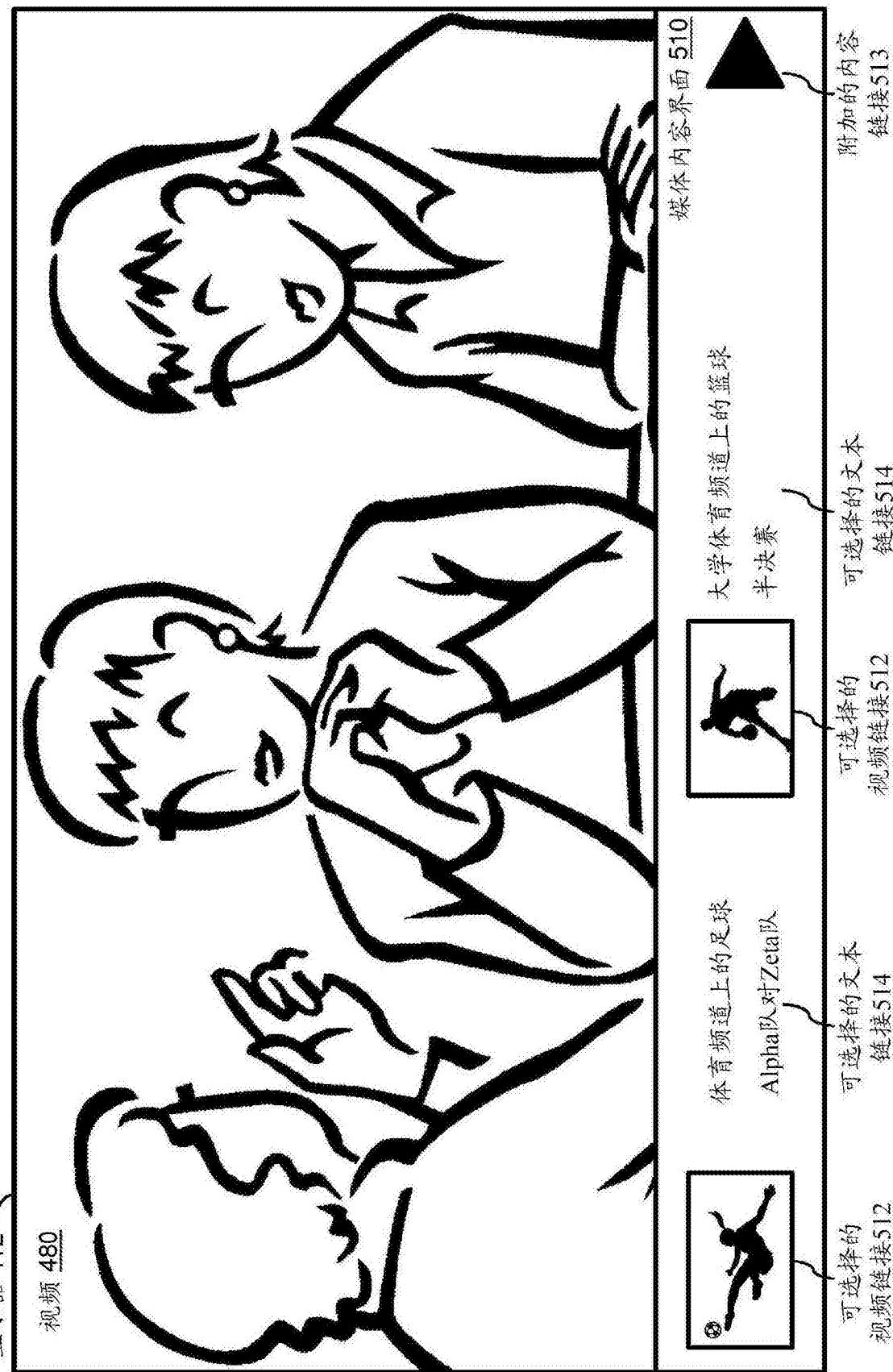


图5

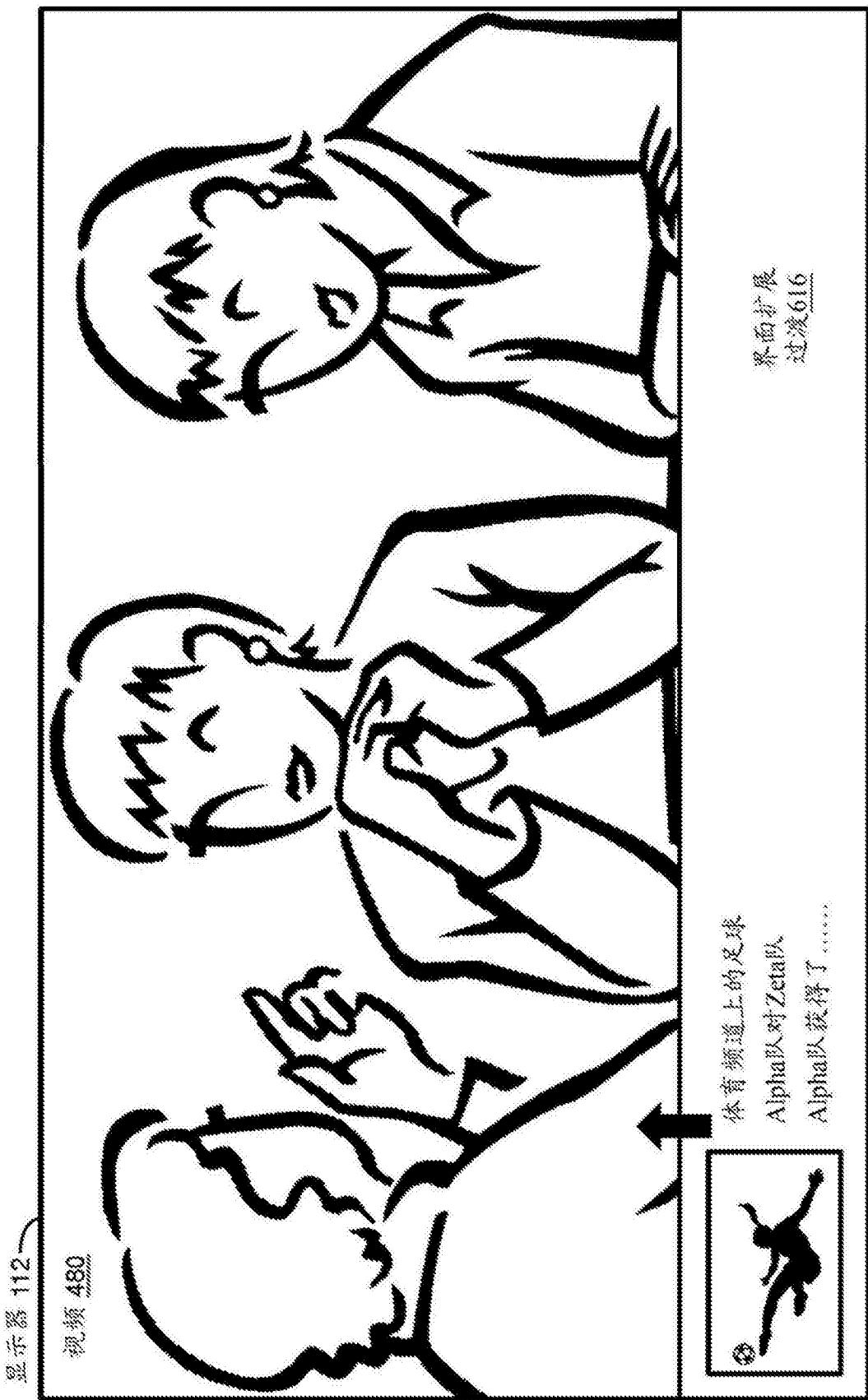


图6A

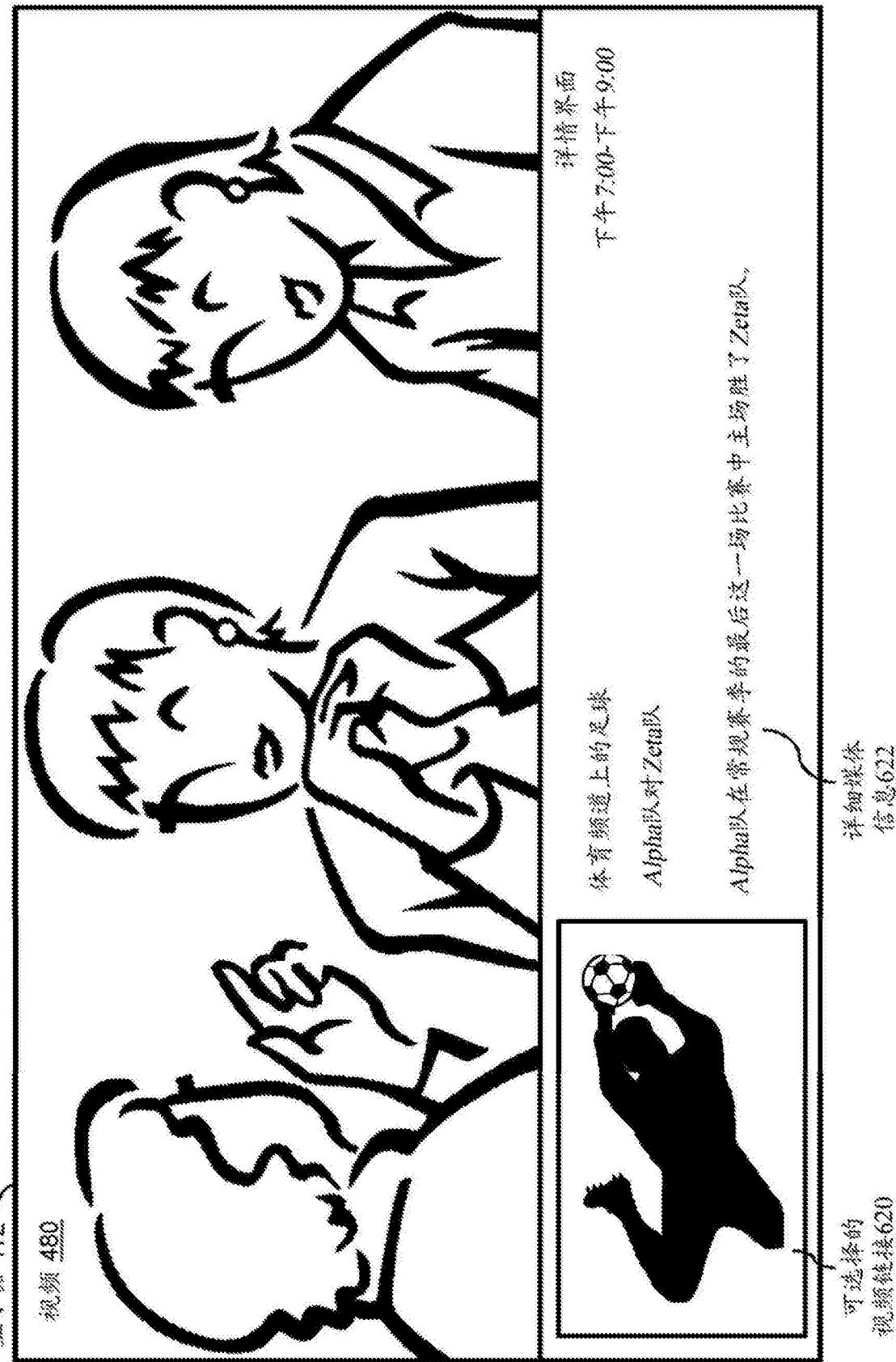


图6B



图7A



图7B

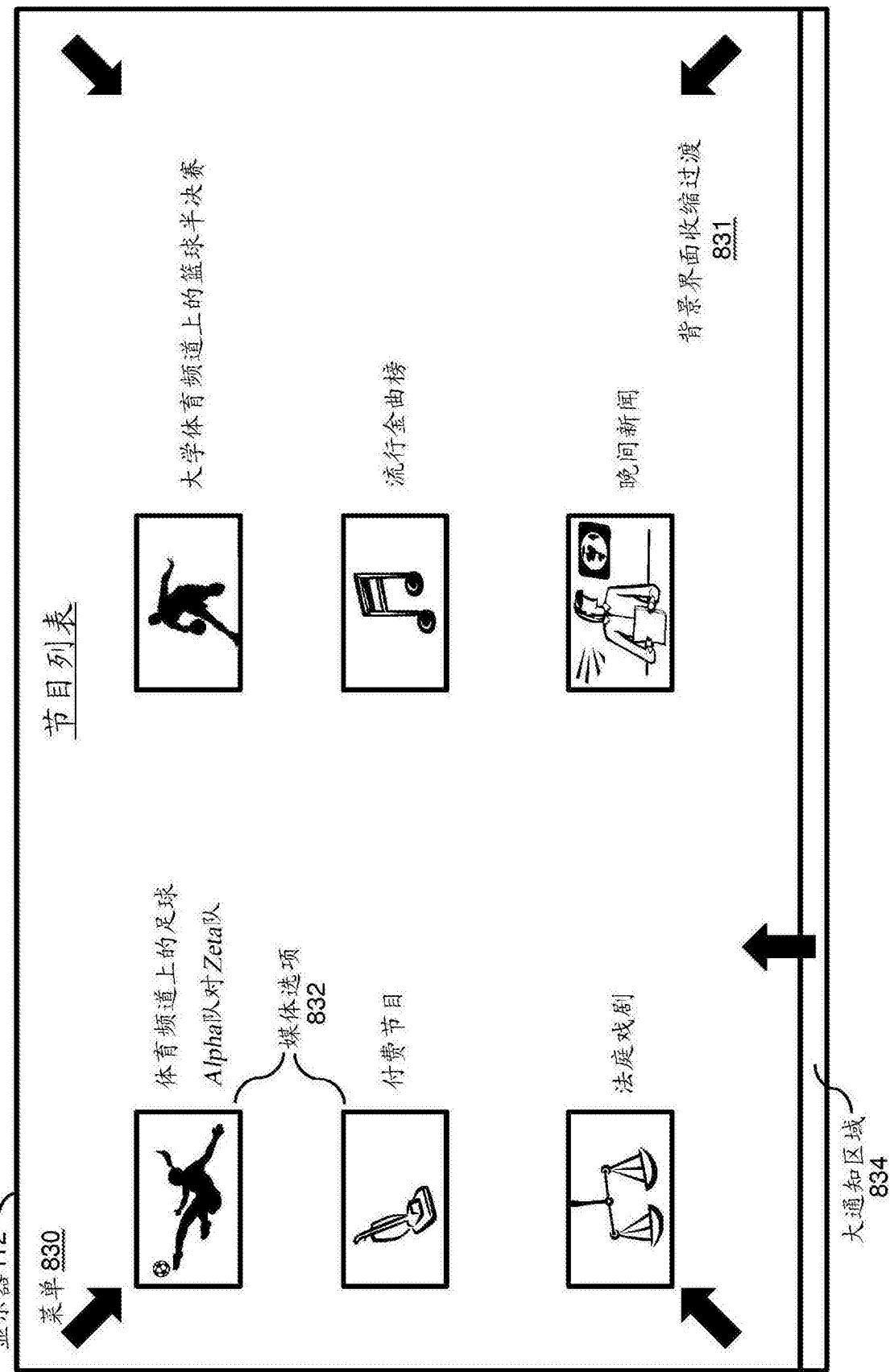


图 8A

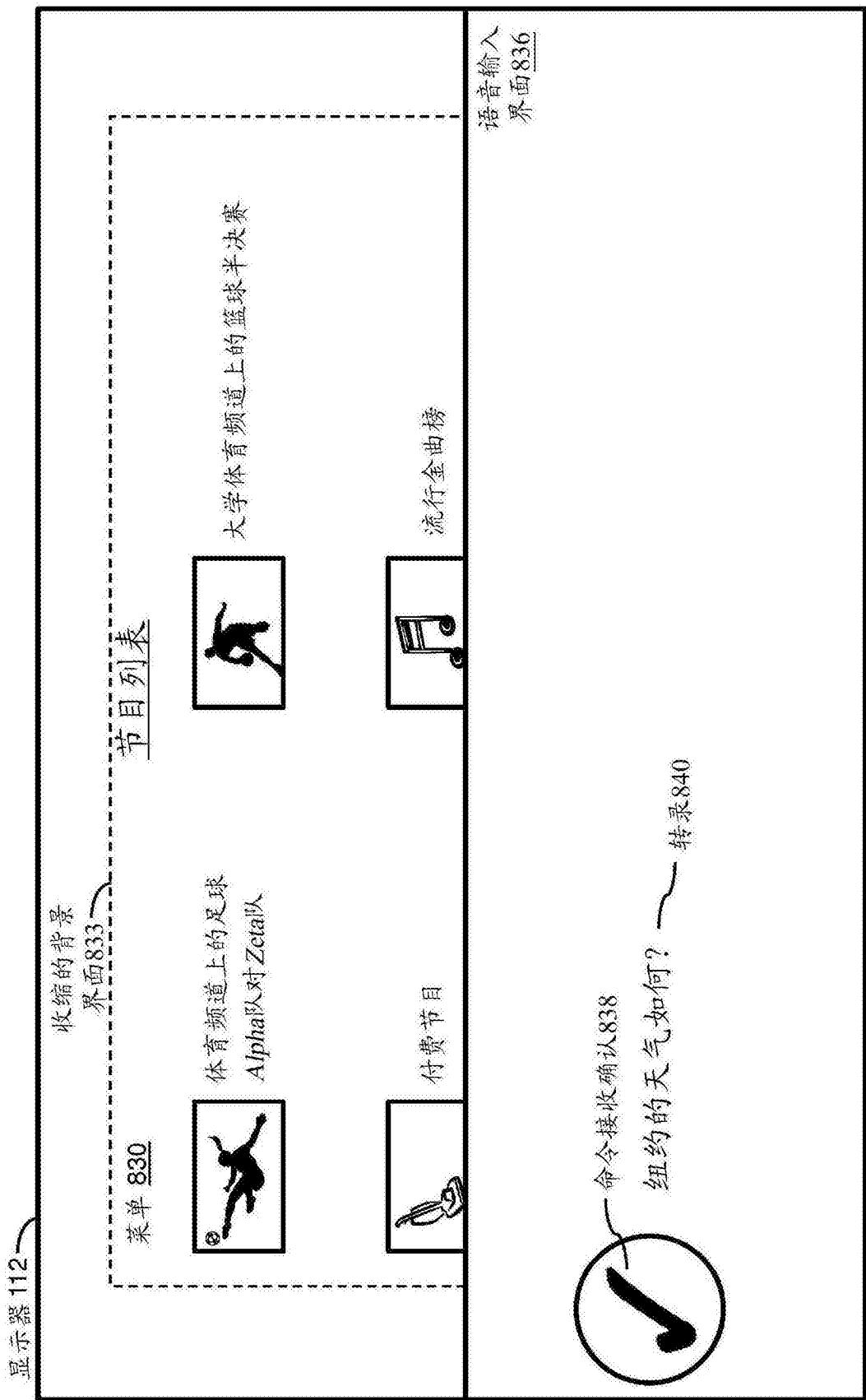
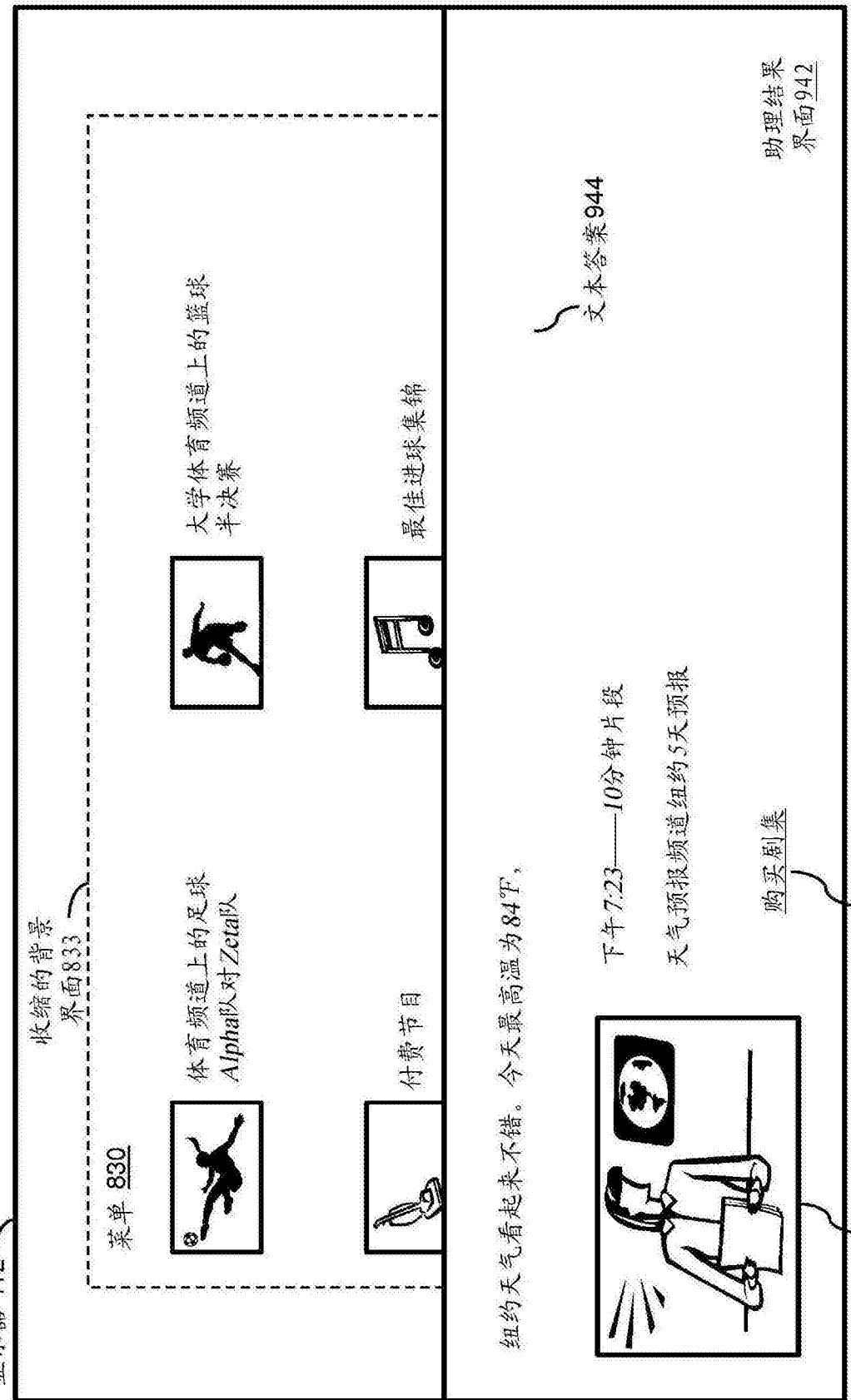


图 8B



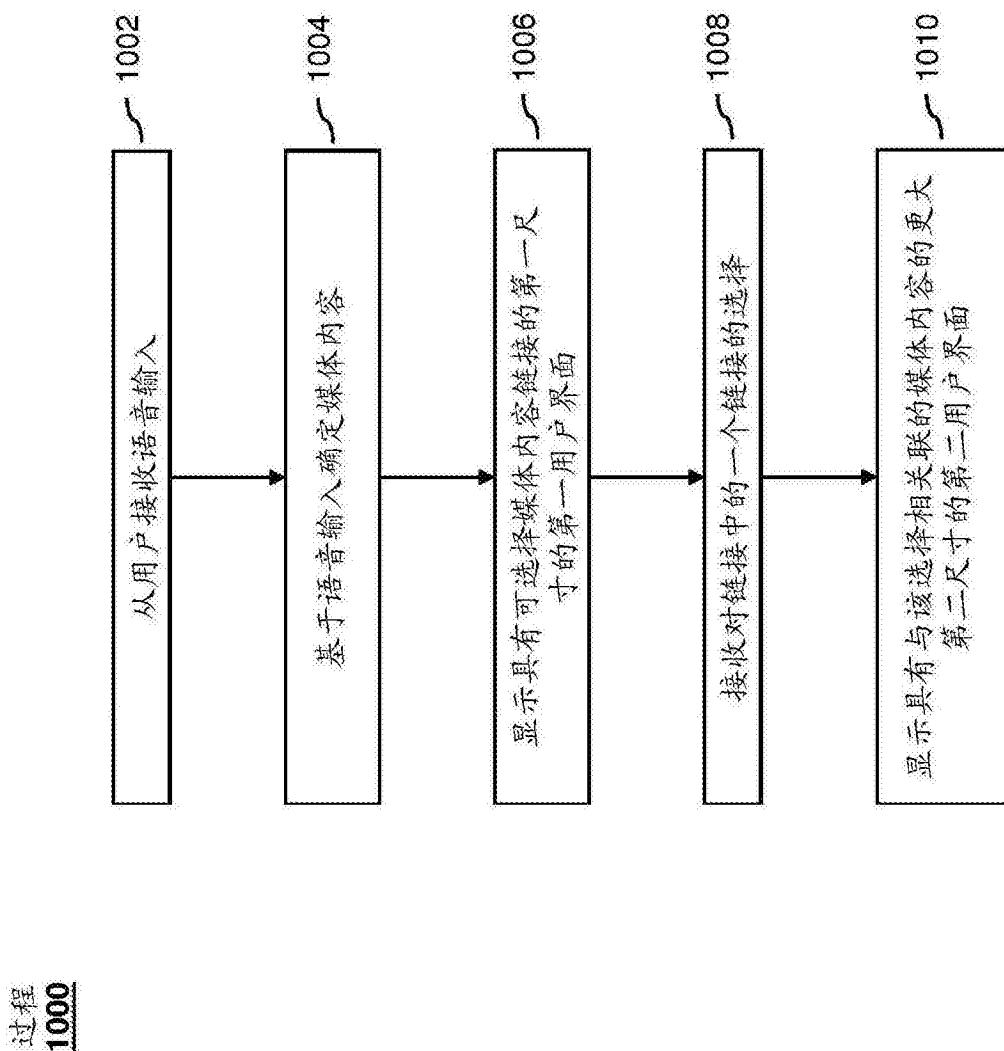


图10

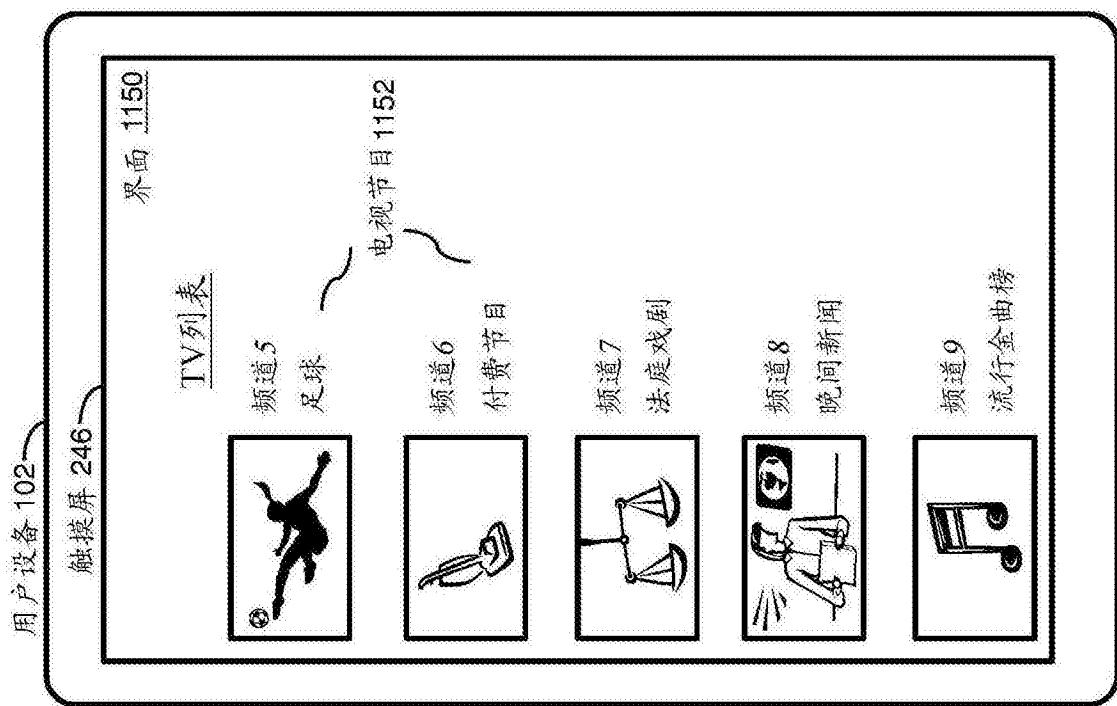


图11

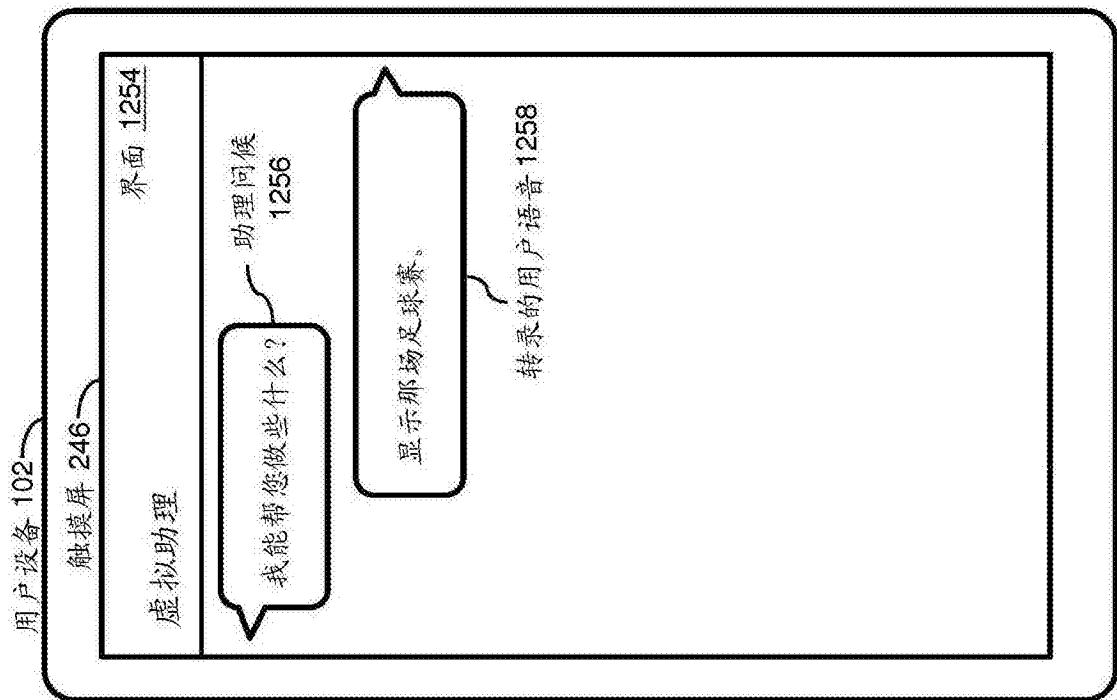
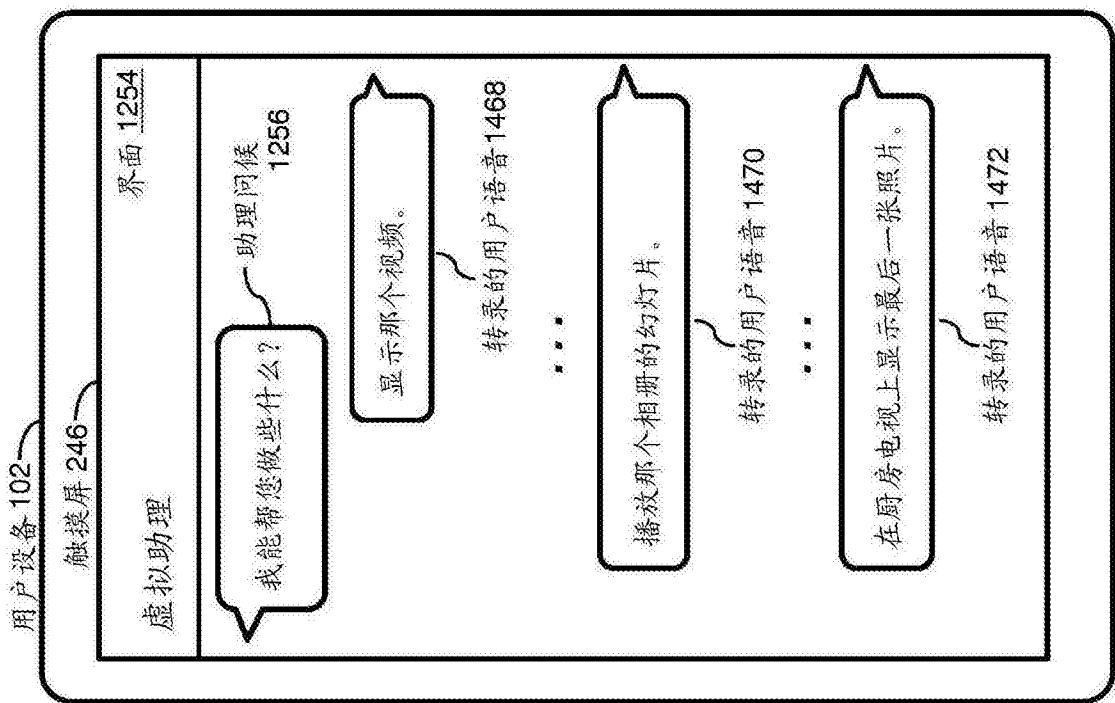
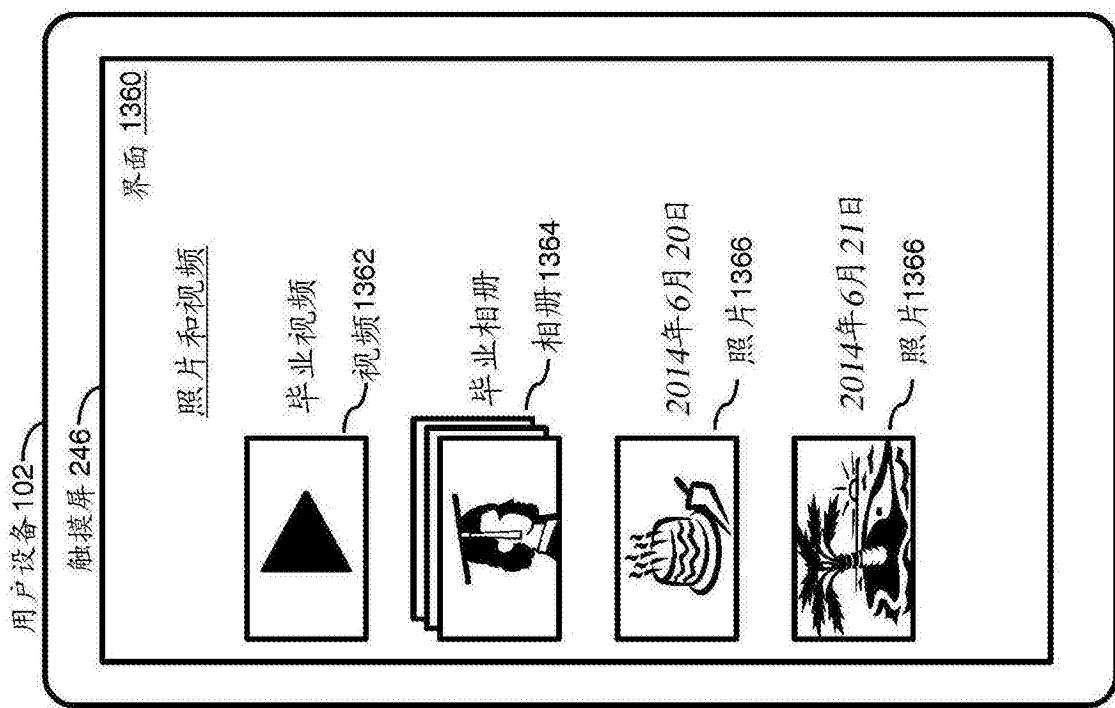


图12



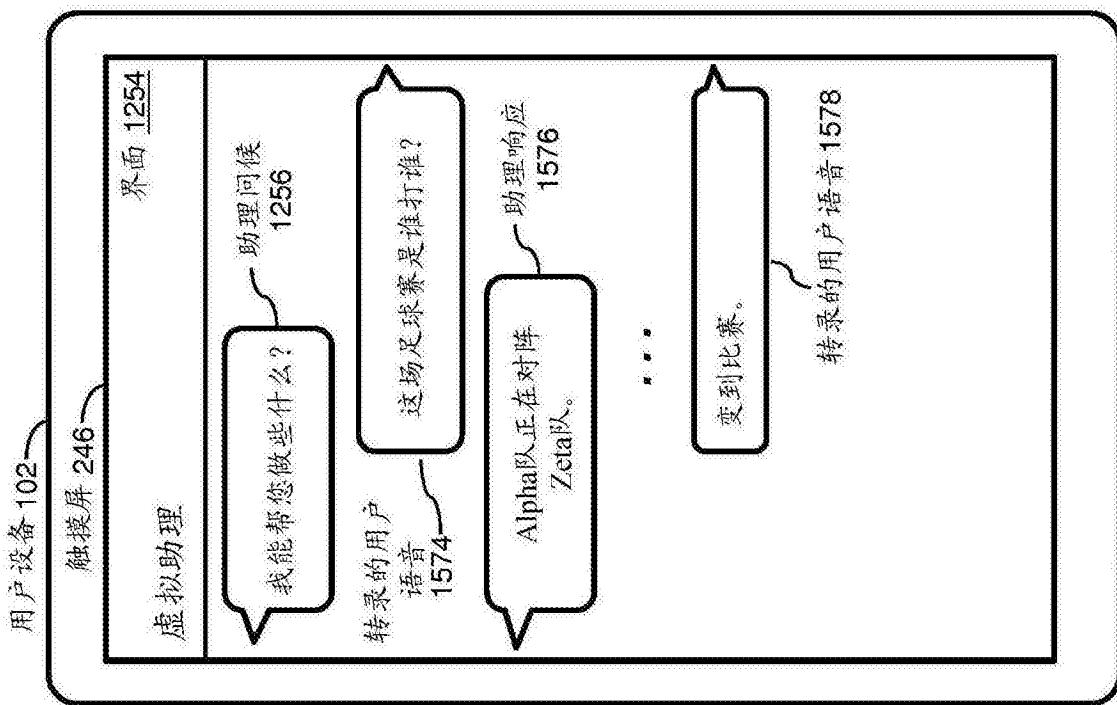


图 15

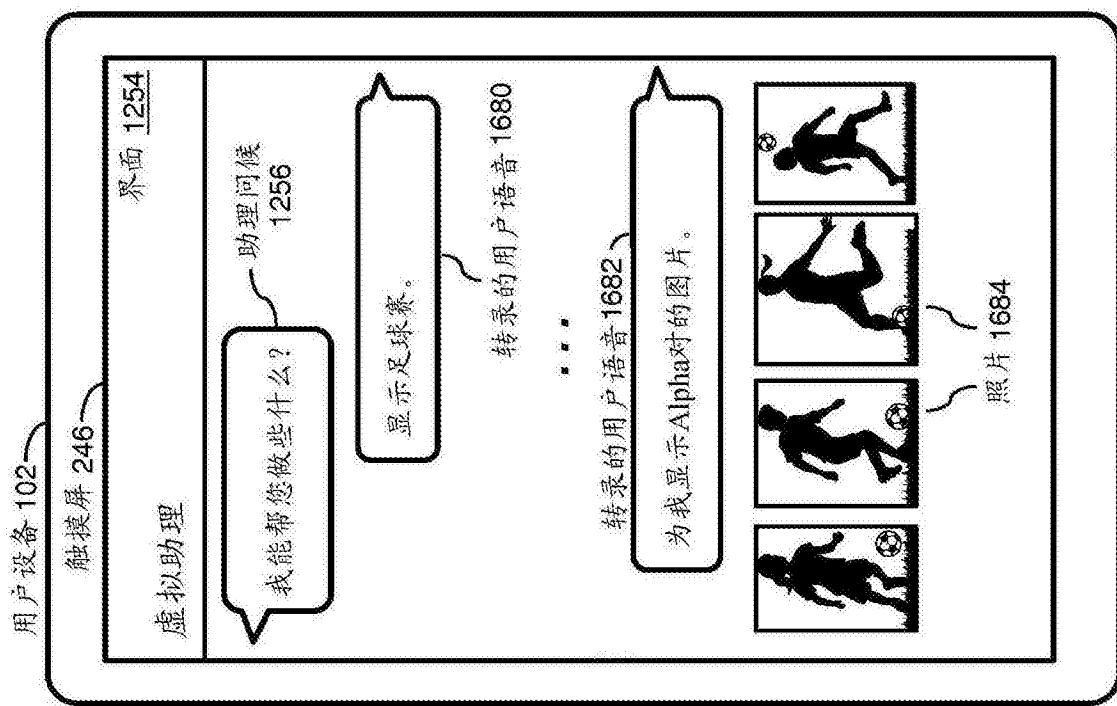


图 16

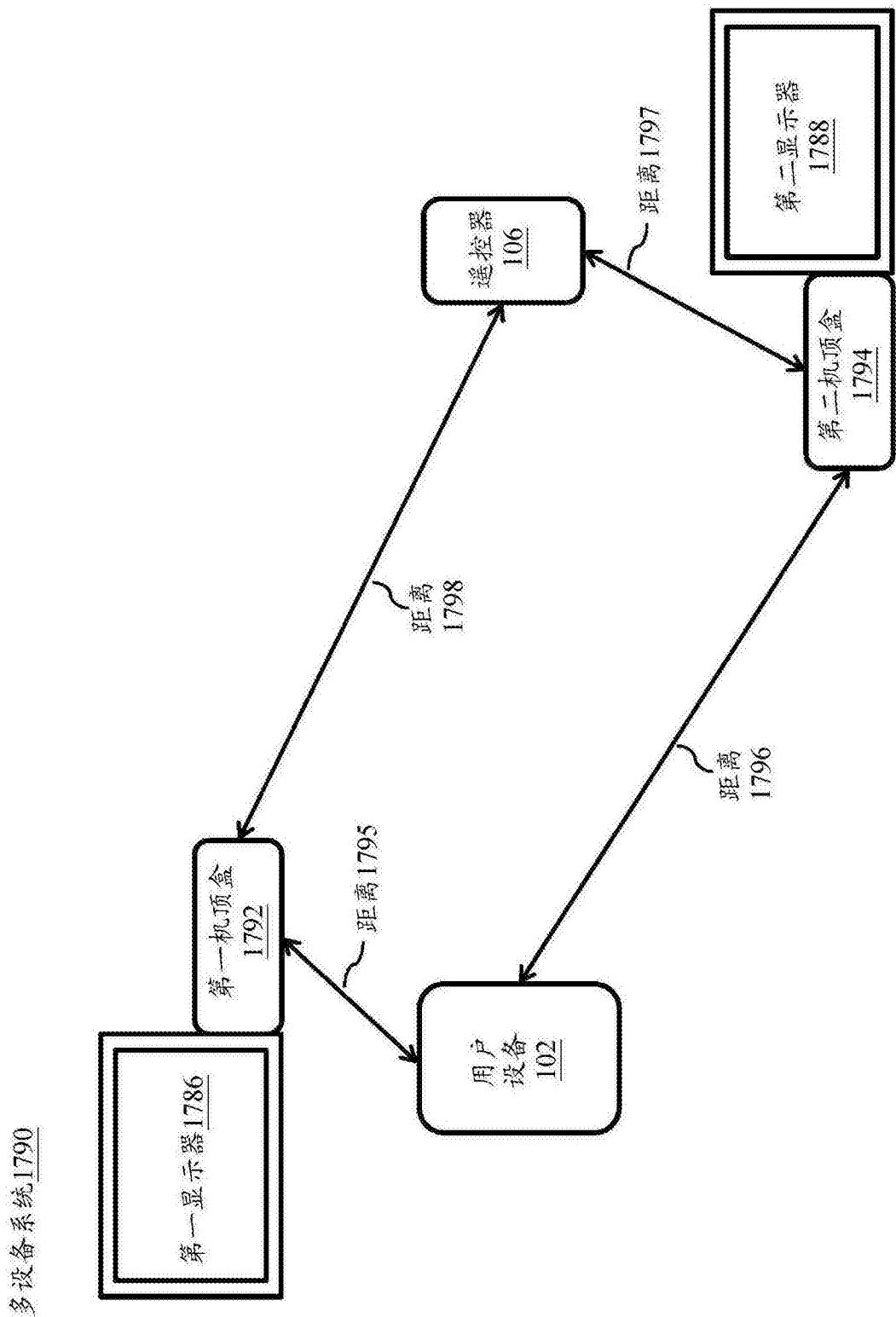


图17

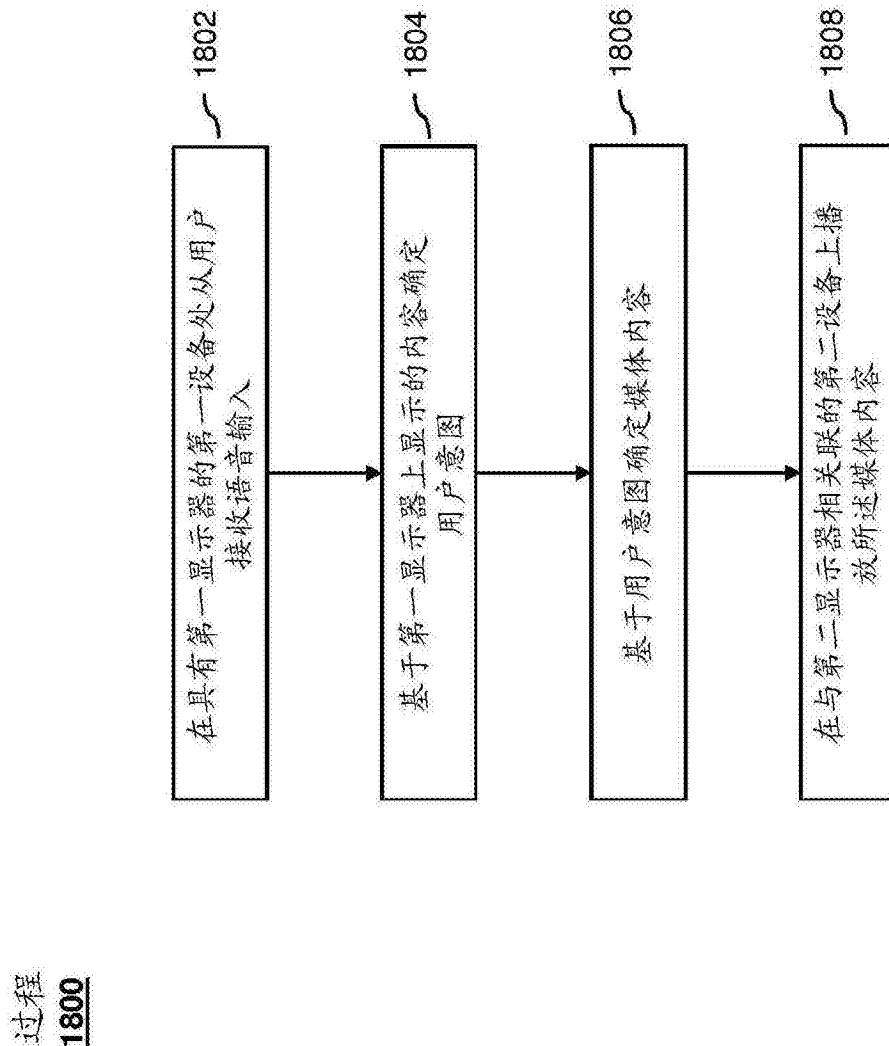


图18



图19

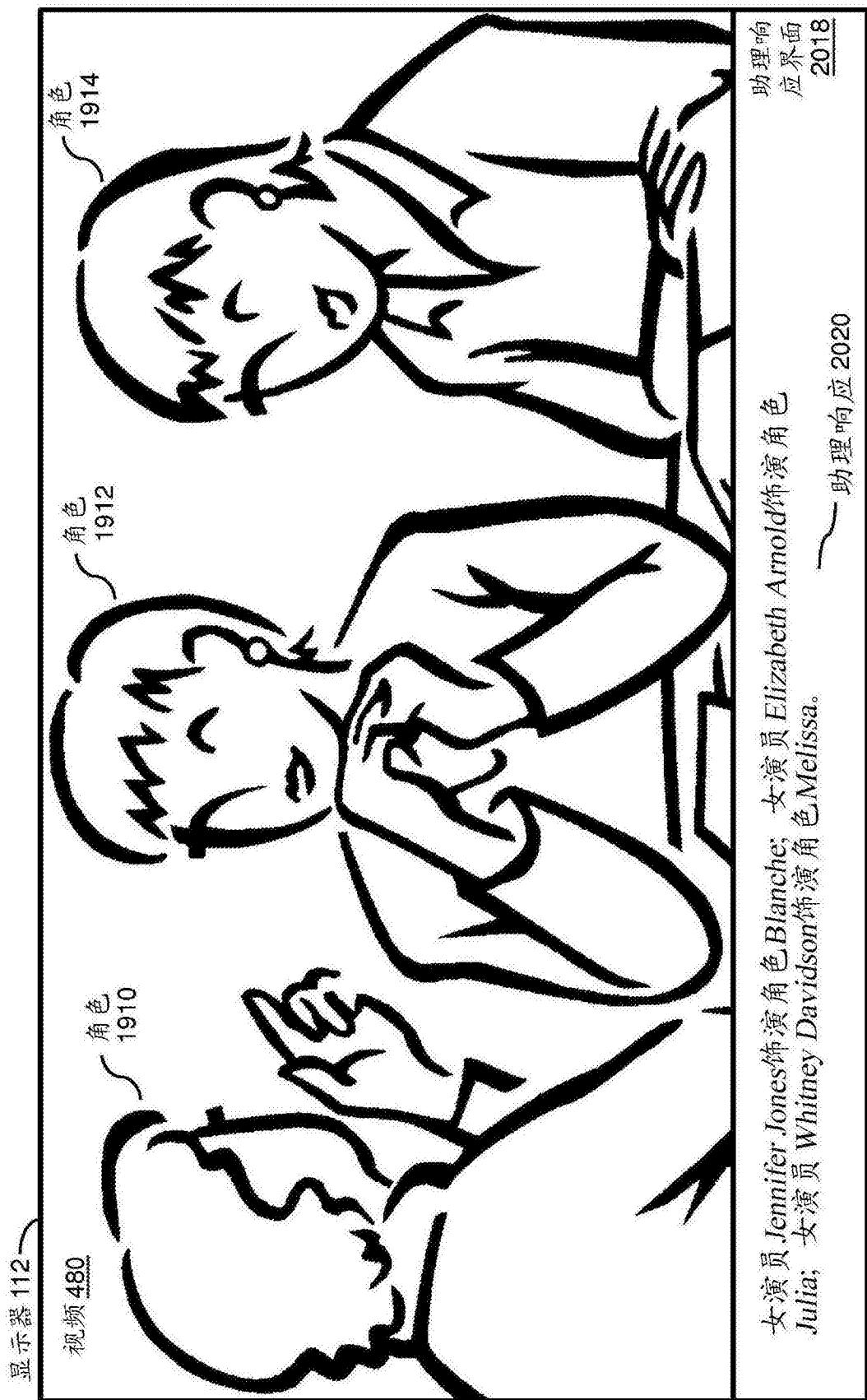


图 20

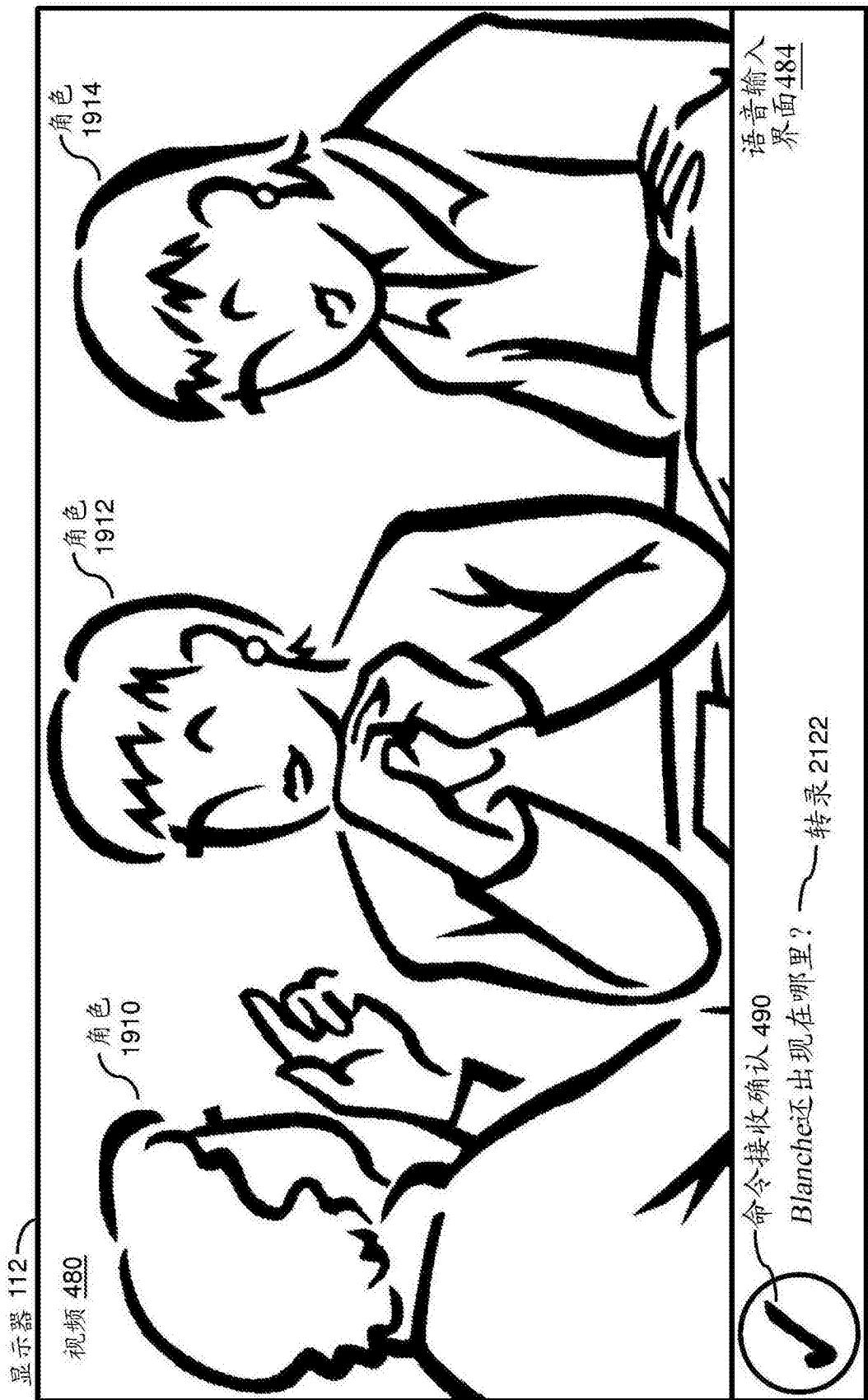


图21

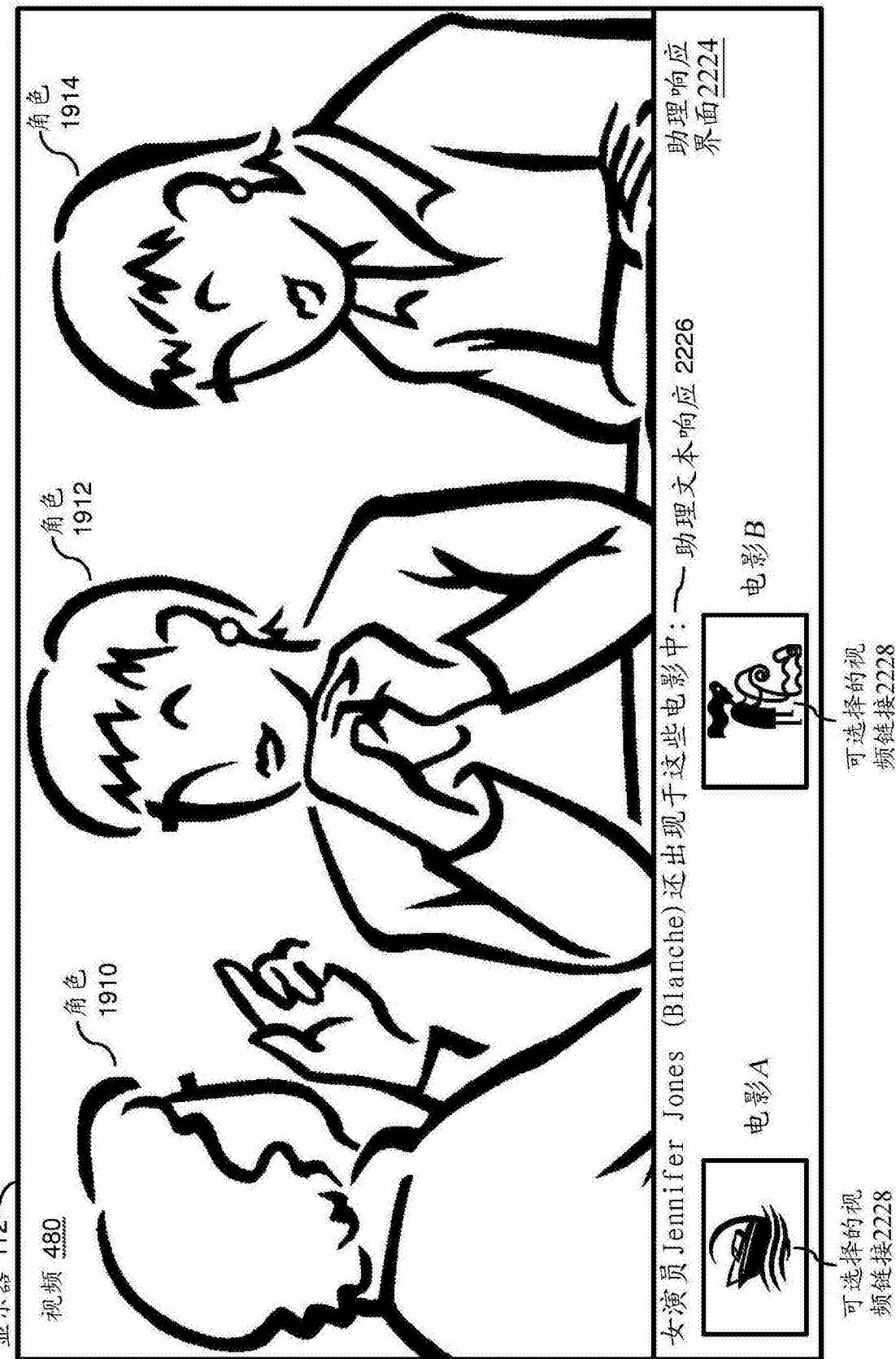


图22

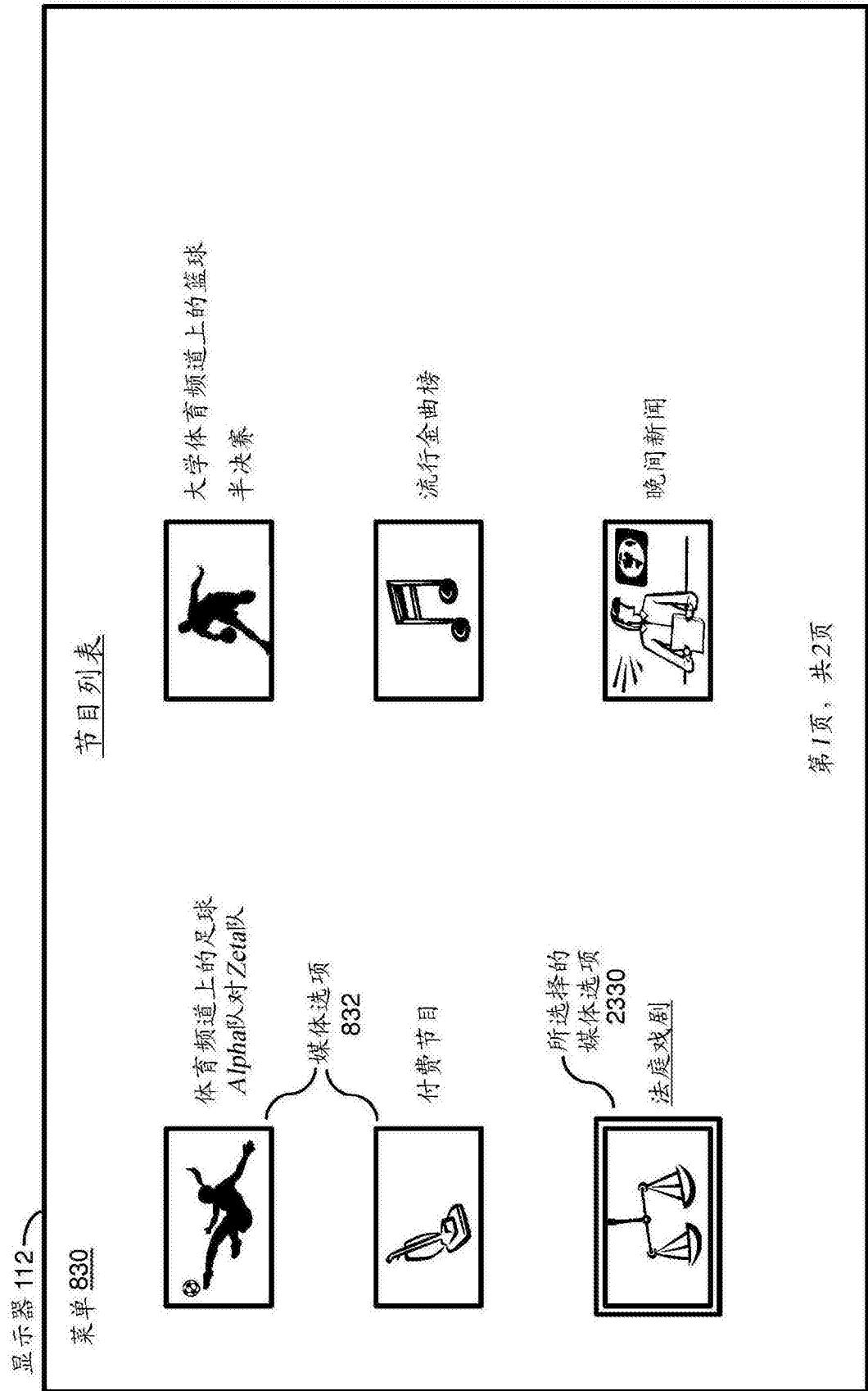


图23A

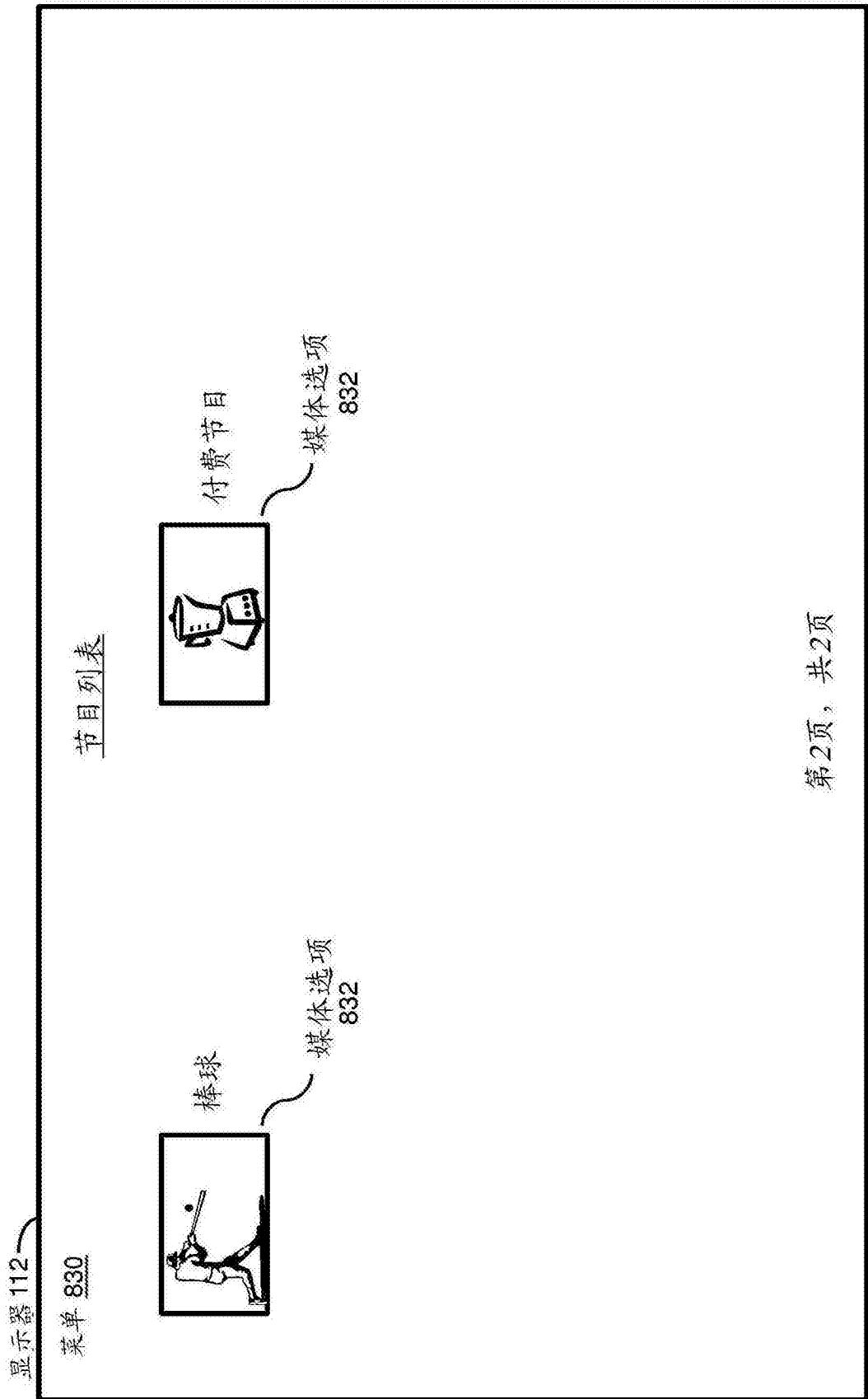


图23B

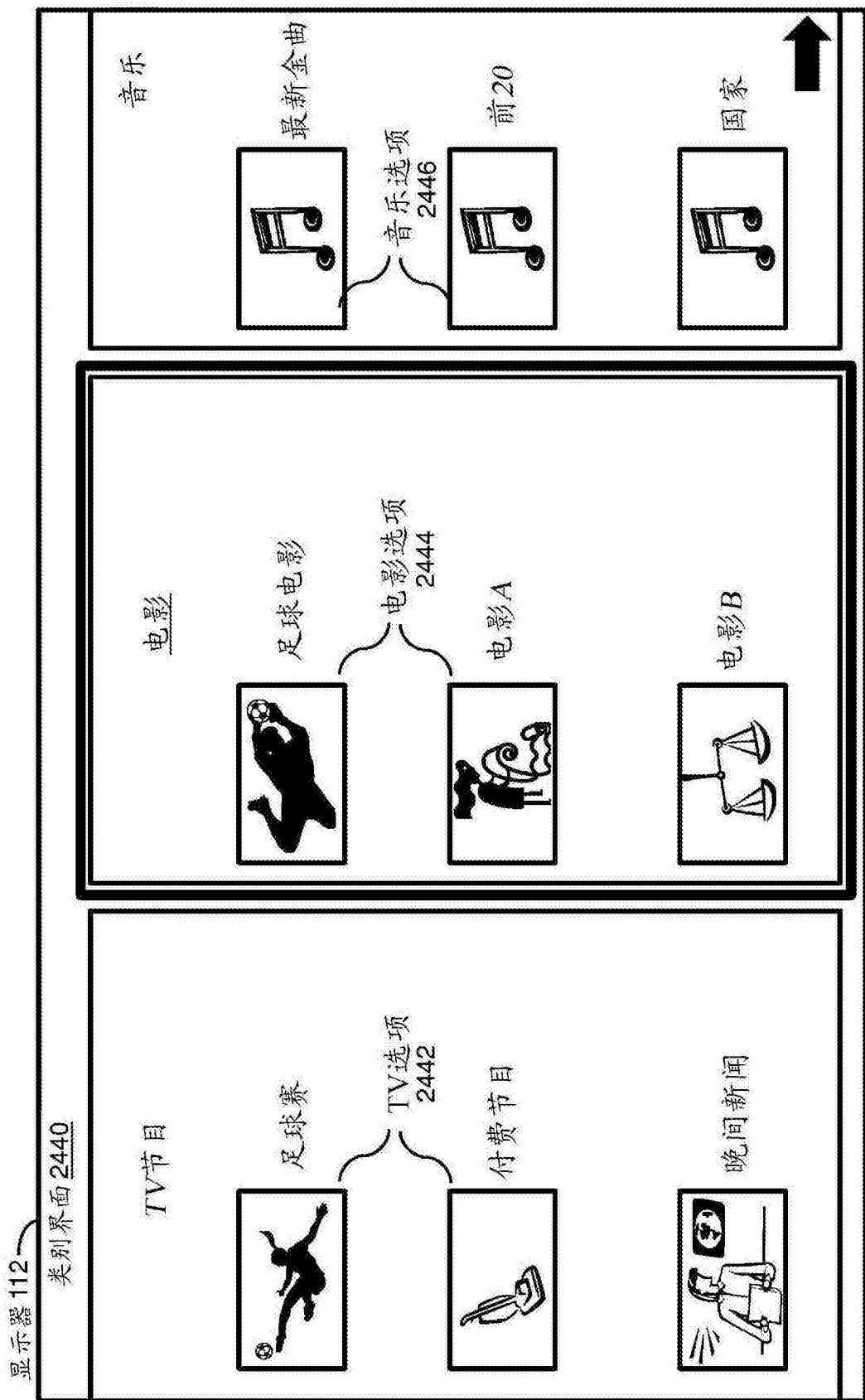


图 24

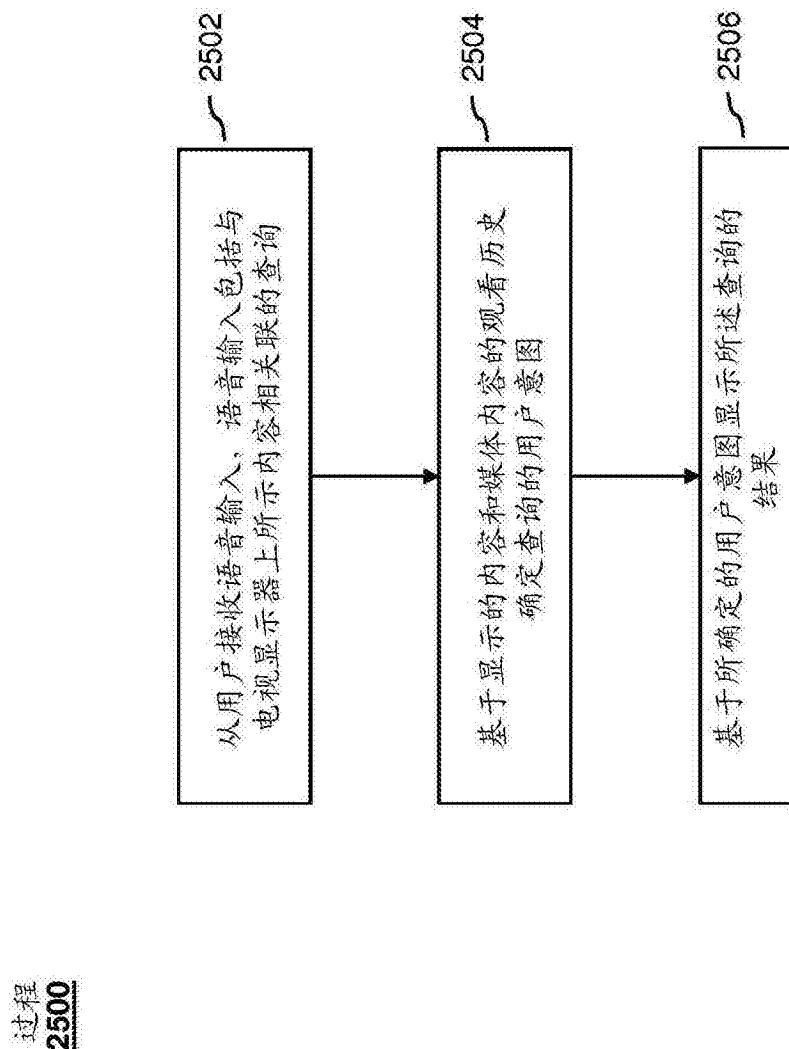


图25



图 26



图27



图28A



图 28B

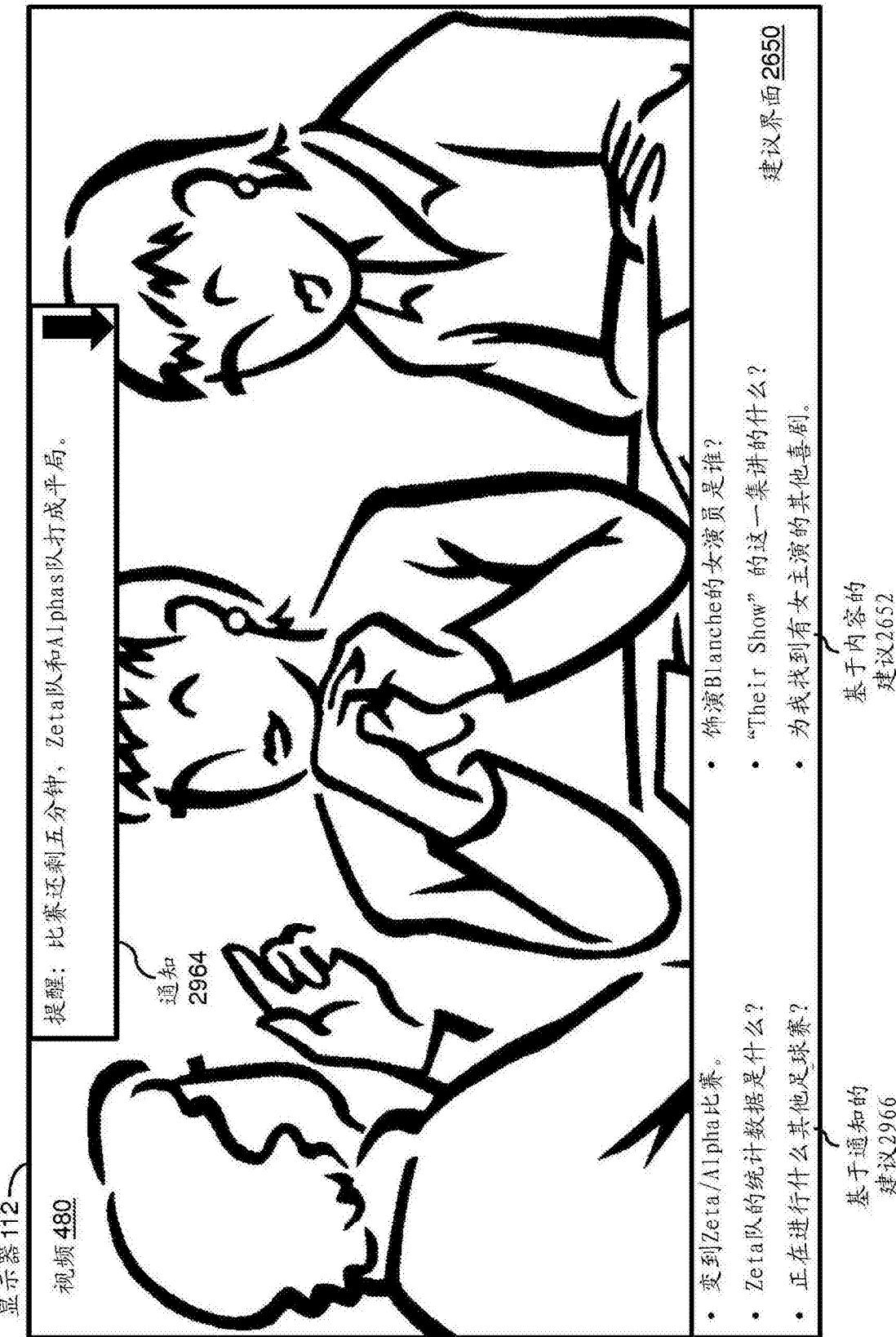


图 29

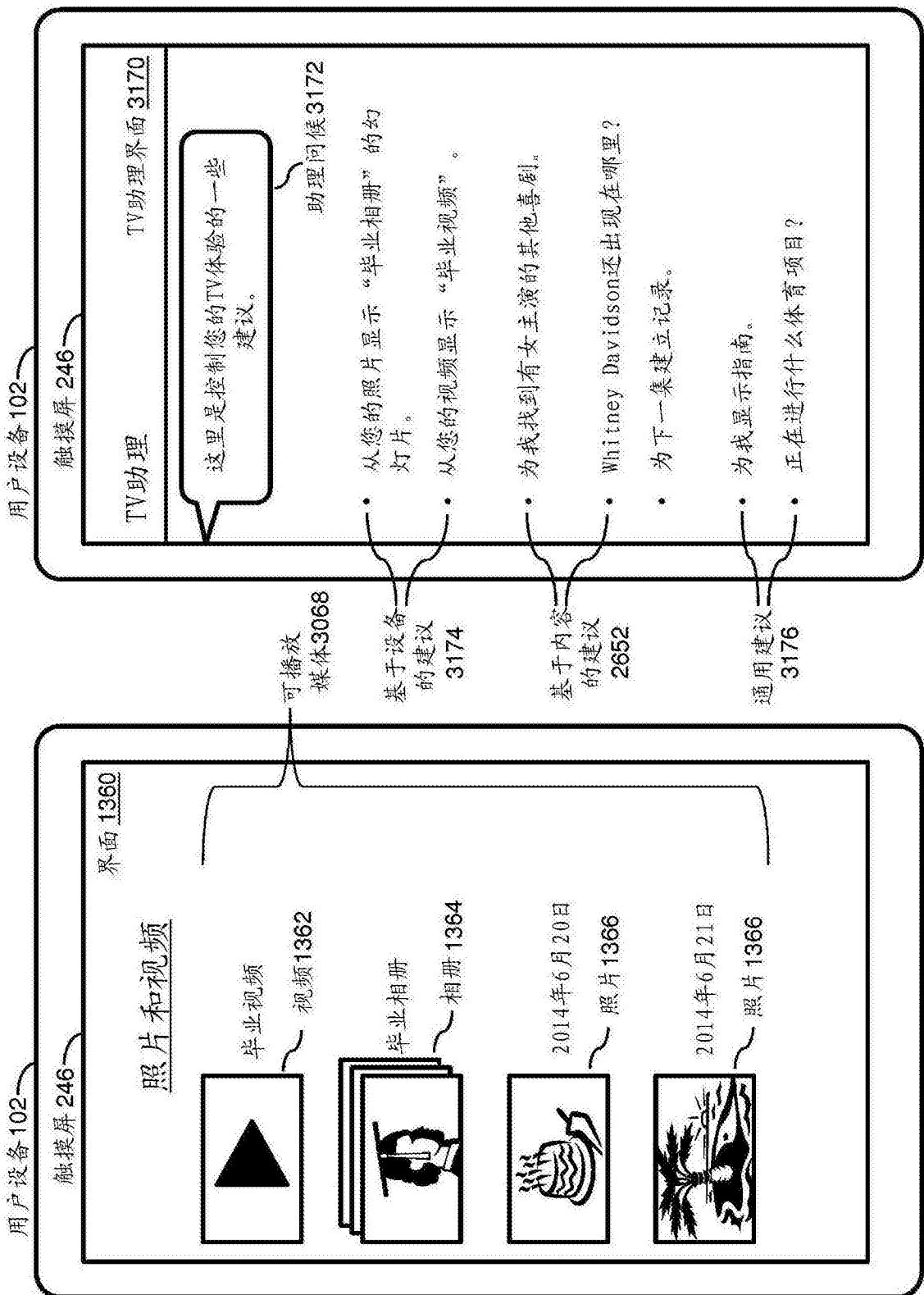


图30

图31

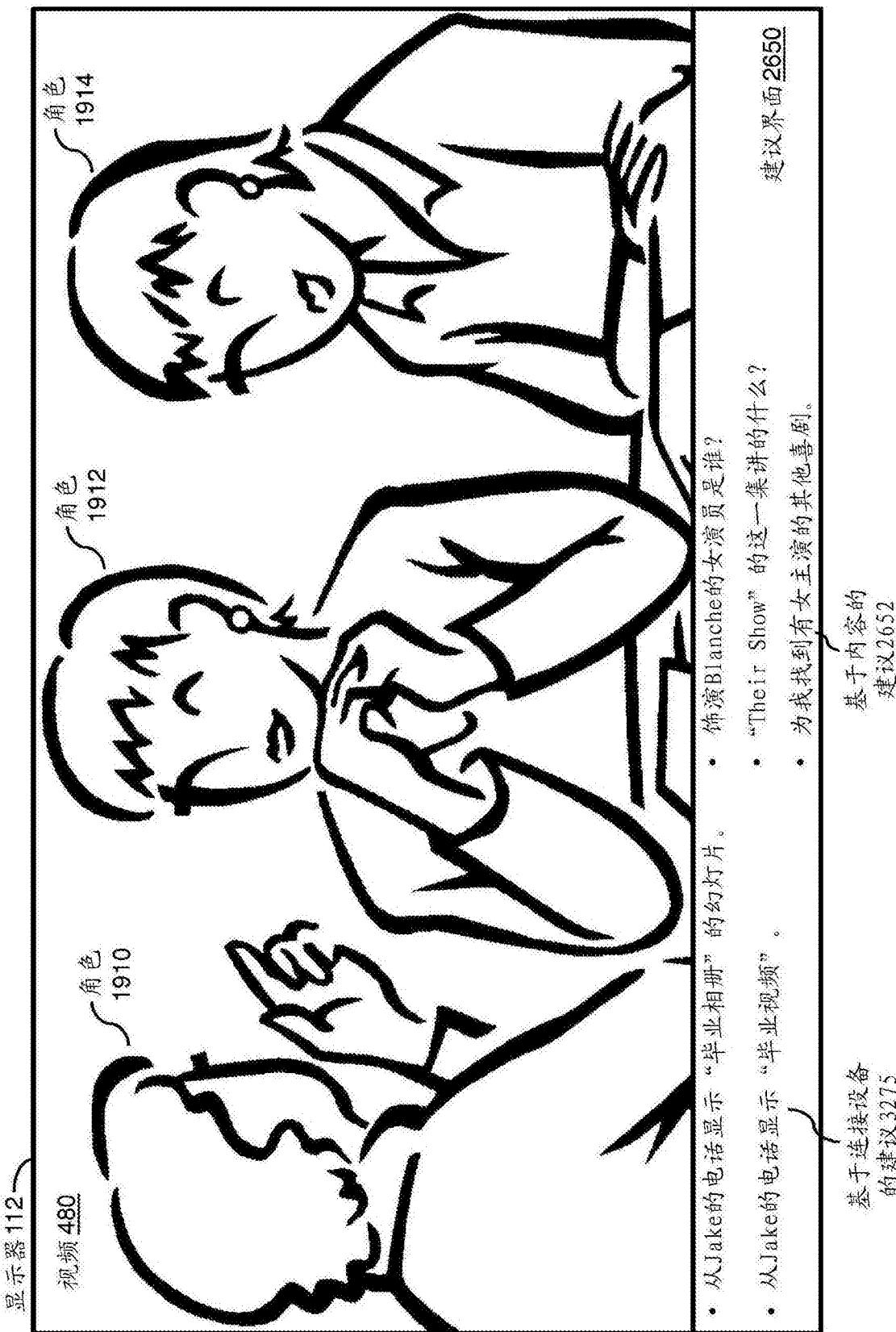


图32

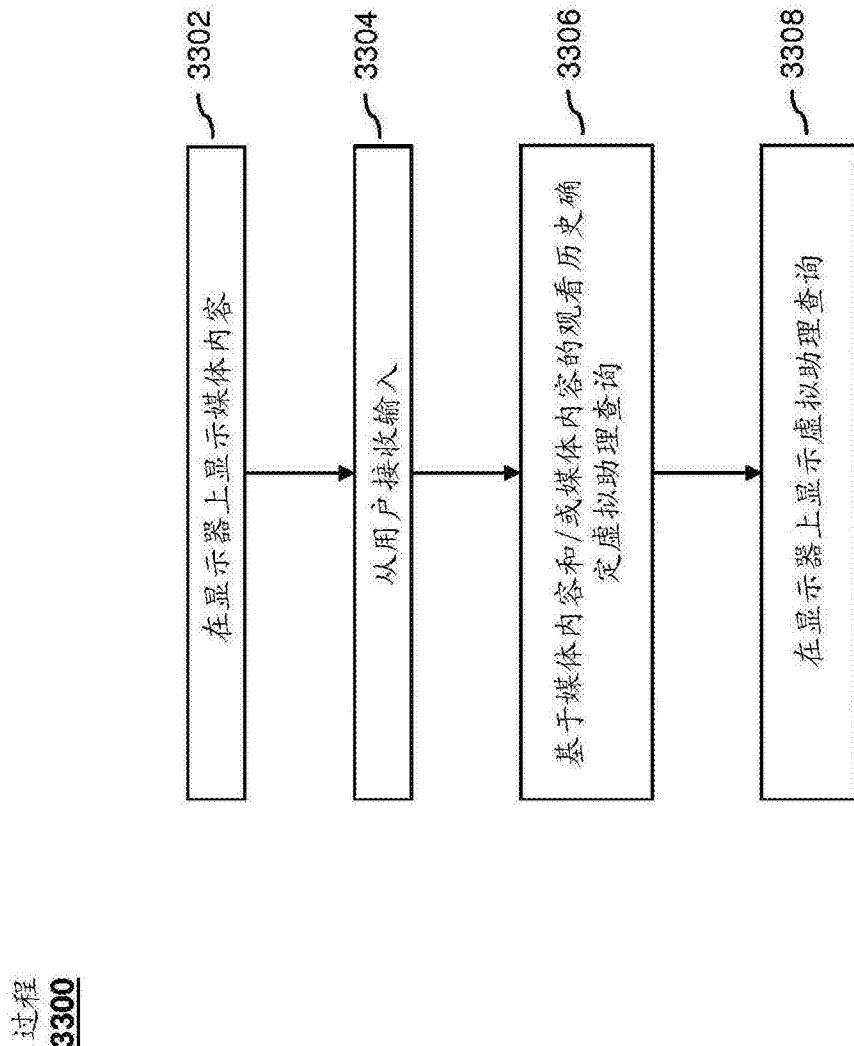


图33

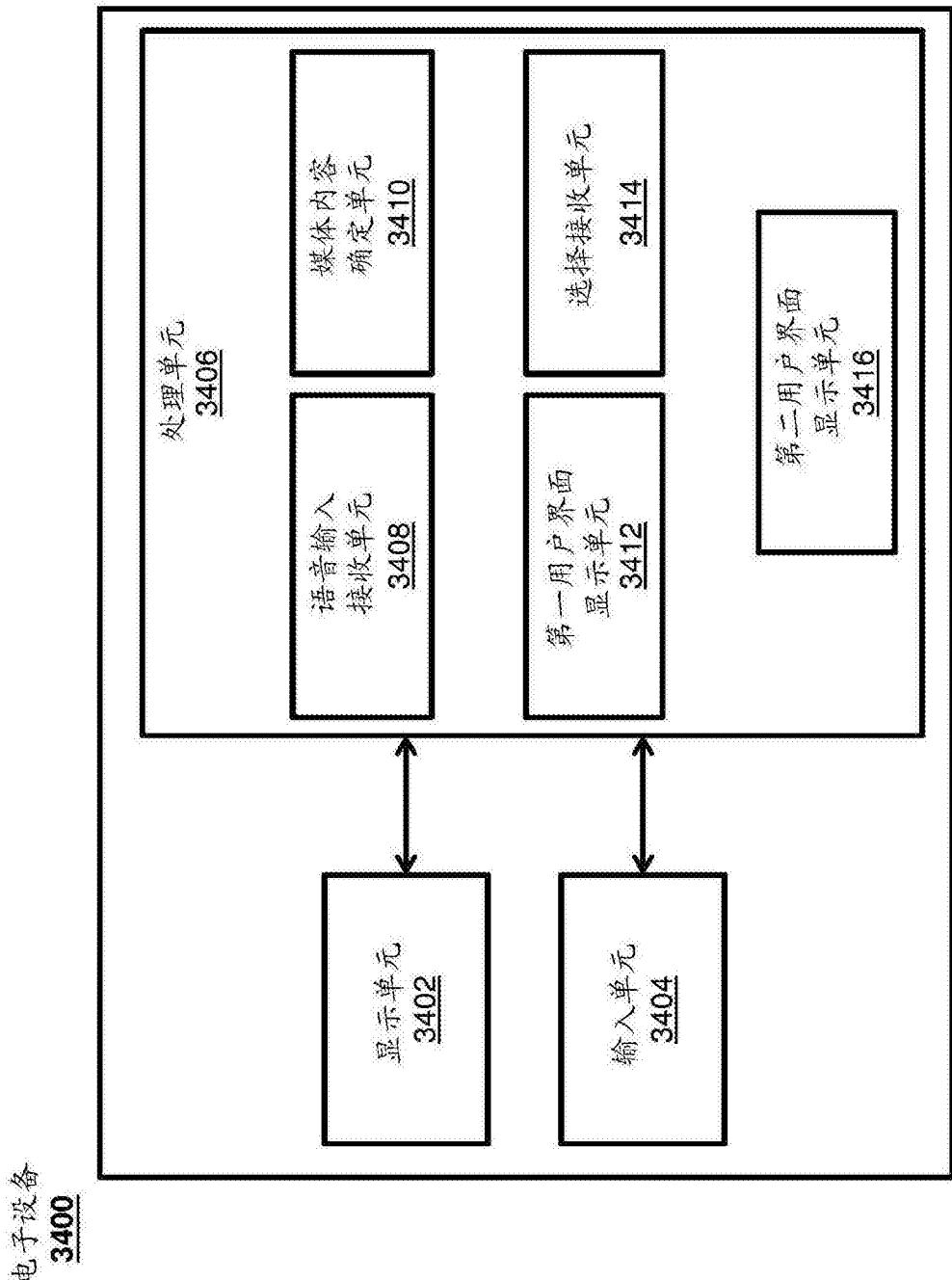


图34

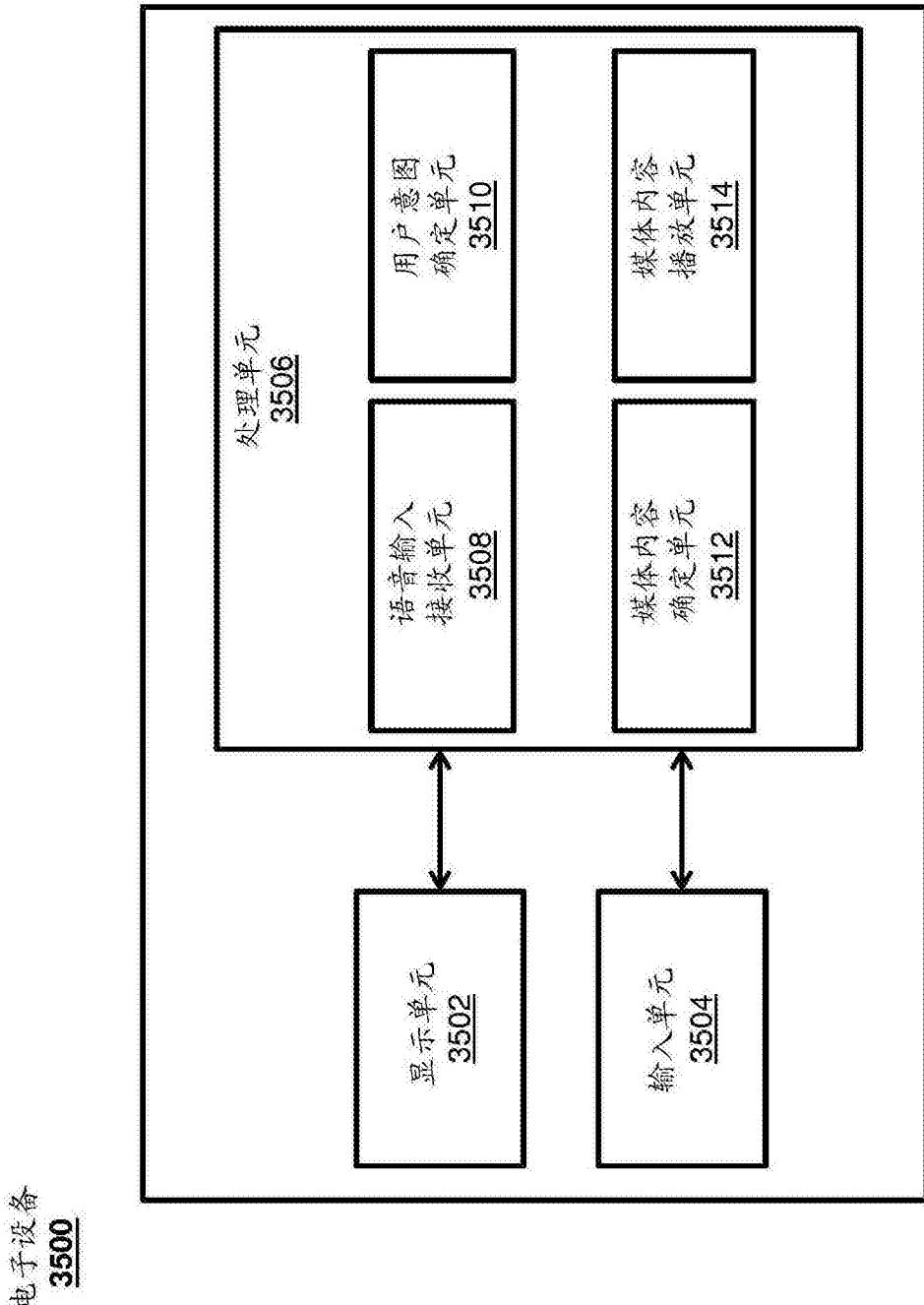


图35

电子设备
3600

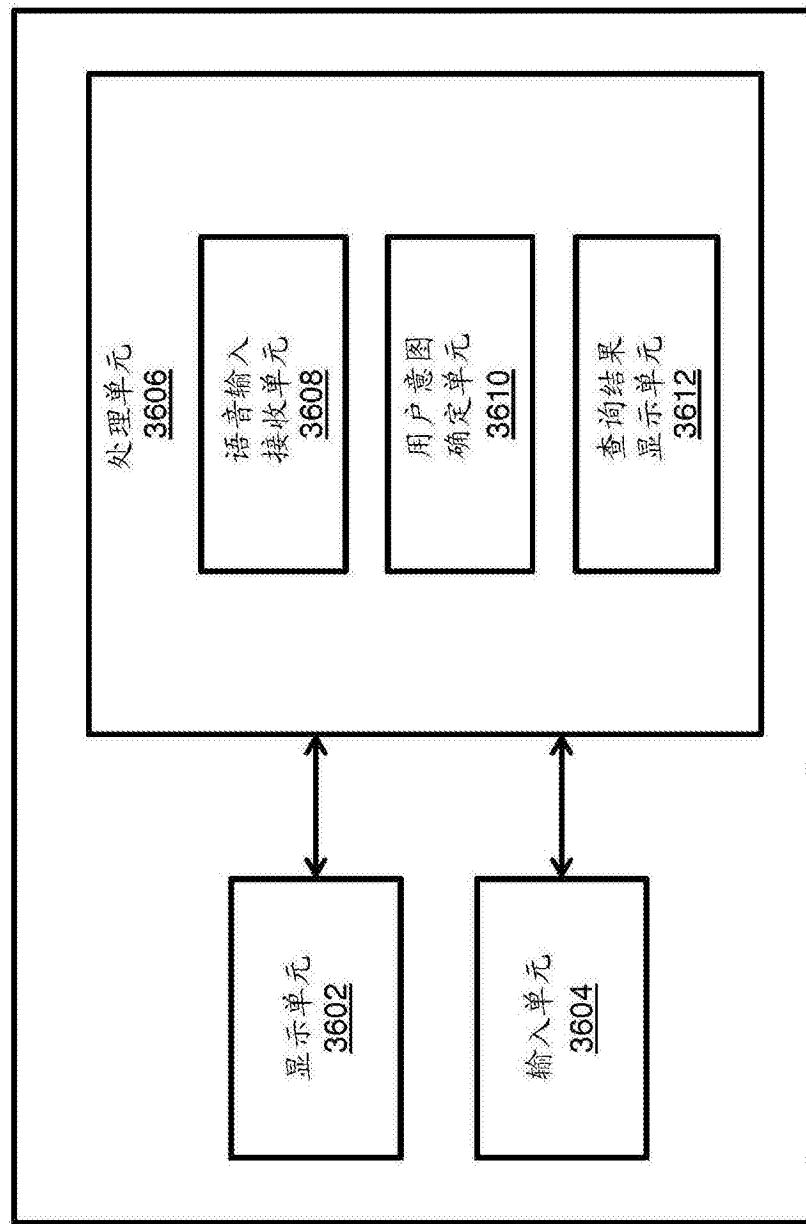


图36

电子设备
3700

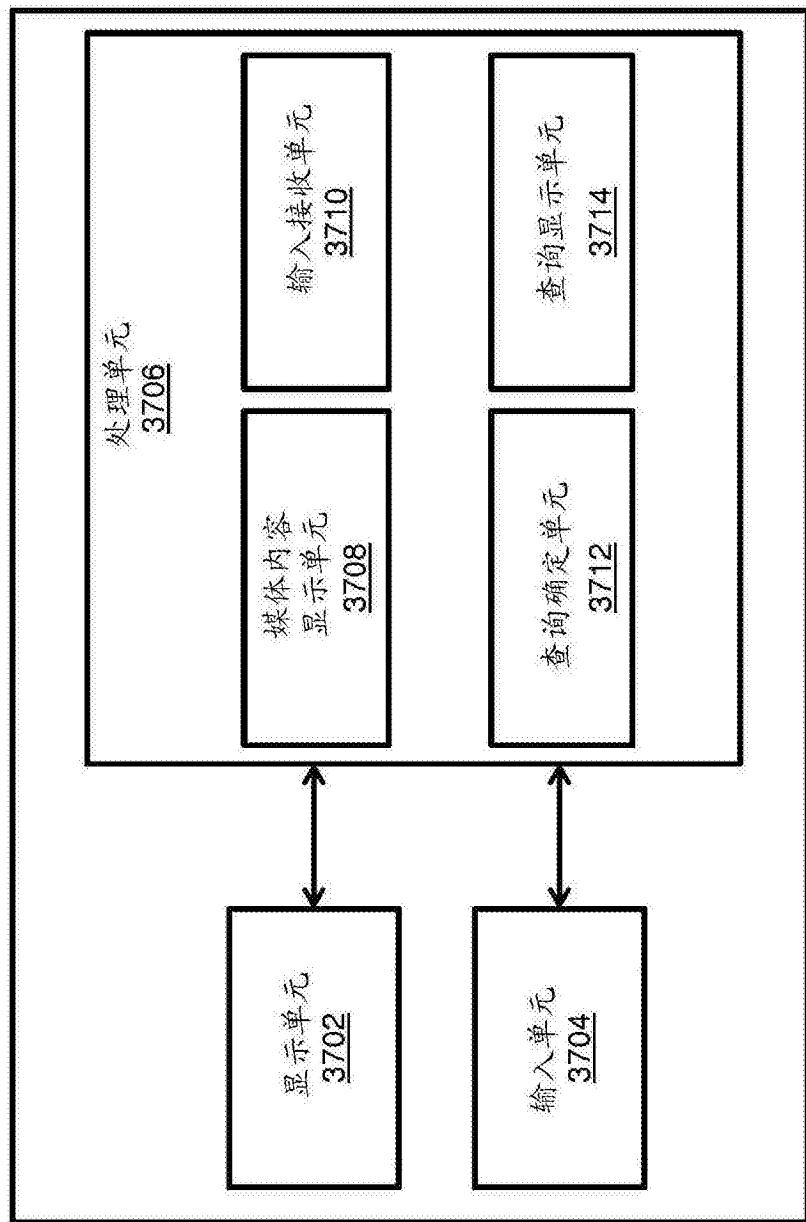


图37